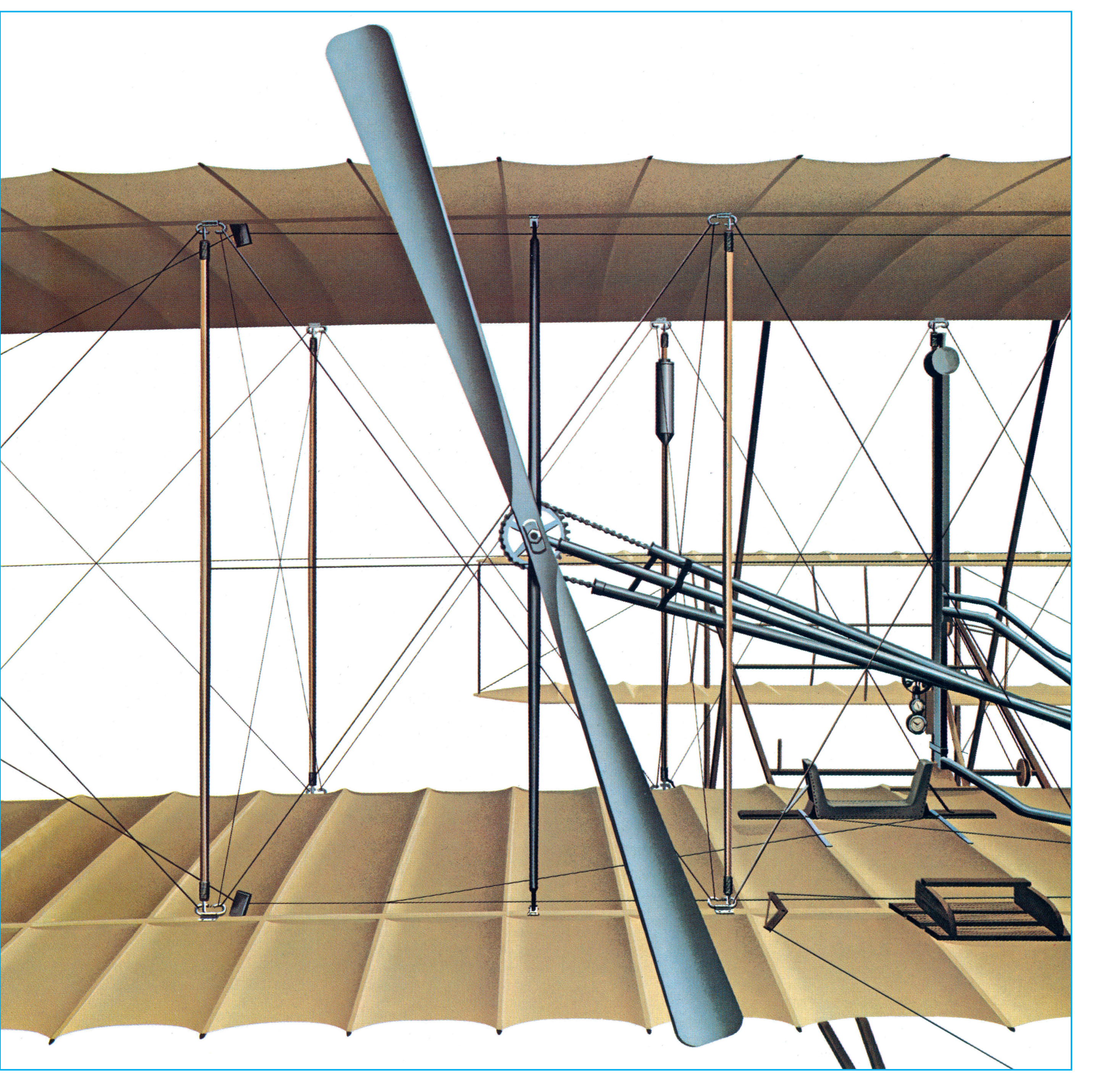
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



AVION MOTOPROPULSADO DE LOS WRIGHT

Septiembre 1979

200 Ptas.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 36
Septiembre 1979

INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN

INVESTIGACION Y CIENCIA

DIRECTOR: Francisco Gracia Guillén
REDACCION: José María Valderas Gallardo (Redactor Jefe)
Encarna de Blas (Secretaria de Redacción)
PRODUCCION: César Redondo Zayas
PROMOCION Y PUBLICIDAD: Elena Sánchez-Fabrés
PROMOCION EXTERIOR: Pedro Clotas Cierco

EDITA: Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

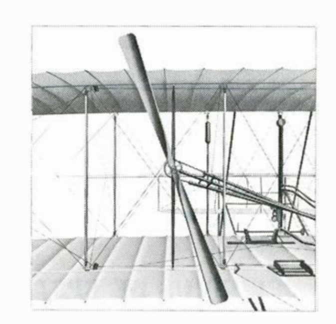
Colaboradores de este número:

Asesoramiento y traducción:

José Luis Somalo: *Los efectos inmediatos y retardados de la guerra nuclear;* Miguel Viñas: *El virus de Epstein-Barr;* Manuel Puigcerver: *El parsec central de la galaxia;* Juan José Martinez: *Los orígenes del primer avión motopropulsado con piloto a bordo;* Ginés Morata: *Compartimentos en el desarrollo de los animales;* Ramón Pascual: *El modelo en bolsa del confinamiento de los quarks;* Joandomènec Ros: *Los animales de Burgess Shale;* Luis Bou: *Juegos matemáticos;* José E. Myro y Alicia Menéndez: *Taller y laboratorio.*

Libros:

Antonio Escarré, Alfredo Garcia González y Luis Mas



LA PORTADA

La ilustración de la portada representa parte del primer avión motopropulsado con piloto a bordo: el avión de 1903, diseñado y construido por Wilbur y Orville Wright (véase "Los orígenes del primer avión motopropulsado con piloto a bordo", por F. E. C. Culick, en este mismo número). La vista está tomada desde la parte posterior del avión y ligeramente hacia la izquierda del mismo. La hélice de 8,5 pies de diámetro, montada detrás, una de las dos que tiene el avión, fue un factor decisivo en el éxito de los Wright. Tuvieron que diseñar la hélice por sí mismos cuando descubrieron que la información publicada sobre hélices marinas era en su mayor parte irrelevante. Las hélices fueron tan eficientes que los hermanos Wright pudieron hacer volar su avión con un motor de gasolina bastante modesto, el cual, como todos los demás elementos, construyeron por sí mismos.

Suscripciones:
Prensa Científica, S.A.
Calabria 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Condiciones de suscripción:
España:
Un año (12 números): 2200 pesetas
Extranjero:
Un año (12 números): 36 U.S.$
Ejemplar atrasado ordinario:
225 pesetas
Ejemplar atrasado extraordinario:
340 pesetas

Distribución para España:
Distribuciones de Enlace, S.A.
Ausias March, 49: Barcelona-10

Distribución para los restantes países:
Editorial Labor, S.A.
Calabria, 235-239: Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137: Madrid-30
Tel. 430 84 81
Cataluña:
Miguel Munill
Balmes, 191, 2.°, 3.ª: Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Diseño de la portada de Ted Lodigensky

Página	Fuente	Página	Fuente
7-8	Alan D. Iselin	69-74	Bunji Tagawa
10-11	Albert Miller	77-85	Walken Graphics
12-13	U.S. Air Force	89	Simon Conway Morris
14-17	Alan D. Iselin	90	Simon Conway Morris (*arriba*), Tom Prentiss (*abajo*)
18-19	Albert Miller	91-93	Tom Prentiss Palaeontological Association, Londres (*arriba y abajo, derecha*) E. Schweizerbart'sche, Stuttgart (*centro y abajo, izquierda*)
23	Werner Henle	95	Albert Miller
24	M. Anthony Epstein, Universidad de Bristol (*arriba*); Ilil Arbel (*abajo*)	96	Royal Society of London (*arriba y abajo*); Palaeontological Association, Londres (*centro*)
25-26	Ilil Arbel	97	Ellis L. Yokelsen (*arriba*), Simon Conway Morris (*abajo*)
27	Gunnar Bjursell, Universidad de Aarhus	98	Simon Conway Morris y H. B. Whittington (*arriba*), Simon Conway Morris (*abajo*)
28-32	Ilil Arbel	101	Universidad de Jena
33	Werner Henle	104-108	Javier de Lorenzo
37	Observatorios Hale (*arriba*); Eric E. Becklin y Gerry Neugebauer, Instituto de Tecnologia de California y Observatorio Hale (*abajo*)	109	Museo de Munich
38-39	George V. Kelvin	110-111	*Scientific American*
40	Observatorios Hale		
41-46	George V. Kelvin		
52-53	Dan Todd		
54-55	Biblioteca del Congreso		
56-64	Dan Todd		
65	Biblioteca del Congreso		
67	Bunji Tagawa		
68	Bunji Tagawa (*arriba*); Peter A. Lawrence, Medical Research Council, Cambridge (*centro y abajo*)		

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición. Tecfa
Guipúzcoa, 36 (local 1) Barcelona-20
Cayfosa. Santa Perpetua de Moguda
Barcelona
Printed in Spain - Impreso en España

Los autores

KEVIN N. LEWIS ("Los efectos inmediatos y retardados de la guerra nuclear") se halla incorporado en el proyecto de control de armamentos en el Centro de Estudios Internacionales del Instituto de Tecnología de Massachusetts, y es consejero de la Rand Corporation. Se graduó en investigación operativa por la Universidad de Yale en 1977, licenciándose en el MIT en la primavera pasada. Las actuales especialidades en investigación de Lewis incluyen la estructuración de las fuerzas armadas convencionales, el impacto económico de la guerra nuclear y la tecnología de la defensa aérea.

WERNER HENLE, GERTRUDE HENLE y EVELYNE T. LENETTE ("El virus de Epstein-Barr") son virólogos del Hospital Infantil de Filadelfia. Los Henle son asimismo profesores de virología pediátrica en la Facultad de Medicina de la Universidad de Pennsylvania. Ambos nacieron en Alemania, y se conocieron de estudiantes en la Universidad de Heidelberg. Por ser Werner nieto de judío, huyó de la Alemania nazi a los Estados Unidos en 1936; Getrude Henle le siguió en 1937. Pertenecen a la plantilla de investigadores del Hospital Infantil y a la Facultad de Medicina de la Universidad de Pennsylvania desde entonces. Este año los Henle han sido galardonados con la segunda edición del premio Bristol Meyers para la investigación sobre el cáncer, en reconocimiento a su trabajo en torno al virus de Epstein-Barr. Lenette nació en China y vivió en Vietnam y Okinawa. Su familia se trasladó a California, donde se graduó en física en la Universidad de California en Berkeley. Se doctoró en la de Washington y se especializó en técnicas virológicas.

THOMAS R. GEBALLE ("El parsec central de la galaxia") es investigador principal en los Observatorios Hale del Instituto de Tecnología de California. Nacido y criado en Seattle se formó en la Universidad de Washington, la de Amsterdam y la de California en Berkeley. Continuó entonces sus estudios en Berkeley, donde obtuvo su doctorado en física en 1974 con una tesis sobre espectroscopía infrarroja aplicada a la astronomía. Después de un año postdoctoral en Berkeley, Geballe pasó otros dos en la Universidad de Leiden, regresando a los Estados Unidos para aceptar un puesto en los Observatorios Hale.

F. E. C. CULICK ("Los orígenes del primer avión motopropulsado con piloto a bordo") es profesor de física aplicada y aerorreactores del Instituto de Tecnología de California. Nacido en New Hampshire, estudió en el Instituto de Tecnología de Massachusetts, antes y después de terminar la carrera, doctorándose en aeronaútica en 1961. Aceptó después una colaboración de postdoctorado en el Cal Tech, donde ha permanecido desde entonces. Sus campos de interés en investigación incluyen lasers de gas, mecánica de fluidos, combustión y el avión de 1903 de los hermanos Wright.

ANTONIO GARCIA-BELLIDO, PETER A. LAWRENCE y GINES MORATA ("Compartimentos en el desarrollo de los animales") son especialistas en biología del desarrollo. García-Bellido es el director del Instituto de Genética del Consejo Superior de Investigaciones Científicas. Obtuvo su doctorado en la Universidad de Madrid en el año 1962 y realizó estudios postdoctorales en las Universidades de Cambridge y Zürich y en el Instituto de Tecnología de California. Lawrence es miembro de la división de biología celular en el Laboratorio de Biología Molecular del Medical Research Council, en Cambridge, Inglaterra. Cursó sus estudios de zoología en Cambridge y obtuvo su doctorado con un trabajo sobre desarrollo de insectos en el laboratorio de V. B. Wigglesworth. De 1965 a 1967 realizó estudios postdoctorales en la Universidad de Virginia y en la de Case Western Reserve. Posteriormente entro en la Facultad de Genética de Cambridge y en 1969 pasó al Laboratorio de Biología Molecular. Morata es miembro del Instituto de Genética de Madrid. Obtuvo su licenciatura en 1968 y entró a trabajar en el laboratorio de García-Bellido al año siguiente donde realizó su tesis doctoral sobre mutaciones que afectan al desarrollo de *Drosophila*. A partir de 1973 trabajo en el departamento de Genética de la Universidad de Oxford y en el laboratorio de Biología Molecular del Medical Research Council en Cambridge, Inglaterra, donde colaboró con Lawrence durante dos años.

KENNETH A. JOHNSON ("El modelo en bolsa del confinamento de los quarks") enseña física en el Instituto de Tecnología de Massachusetts. Hizo sus trabajos de graduación en el Instituto de Tecnología de Illinois y recibió su Ph. D. en la Universidad de Harvard en 1955. Después de dos años profesando en Harvard pasó al Instituto de Física Teórica de Copenhague como becario postdoctoral. En 1958 entró en la facultad del MIT.

SIMON CONWAY MORRIS y H. B. WHITTINGTON ("Los animales de Burgess Shale") son paleontólogos que han estudiado la rica cosecha de fósiles procedente de la localidad de Burgess Shale, en Columbia Británica. Conway Morris es catedrático de geología en la Open University de Milton Keynes, Inglaterra. Nacido en Londres, hizo sus estudios en la Universidad de Bristol y obtuvo su grado de doctor por la Universidad de Cambridge en 1976. En 1975 fue elegido investigador del St. John's College, Cambridge, convirtiéndose en el primer paleontólogo elegido desde L. S. B. Leakey. Whittington es Woodwardian Professor de geología en Cambridge y jefe del departamento de geología. Se educó en la Universidad de Birmingham, donde se doctoró y enseñó de 1945 a 1949. Se incorporó posteriormente a la Facultad de Geología de la Universidad de Harvard, donde hasta 1966 fue conservador de paleontología de invertebrados en el Museum of Comparative Zoology. En 1966 y 1967 formó parte de los equipos del Canadian Geological Survey que recolectaron fósiles en la localidad de Burgess Shale.

JAVIER DE LORENZO ("Frege") es doctor en filosofía y licenciado en matemática por la Universidad de Madrid. Catedrático de matemática del instituto de bachillerato "Zorrilla" de Valladolid. Dedica su atención preferente a la lógica y filosofía de la matemática, temas en los que ha publicado varios ensayos y libros, entre los cuales cabe mencionar, por su enlace con el tema desarrollado en este número, "La filosofía de la matemática de Poincaré" y "La matemática y el problema de su historia".

Hace...

José M.ª López Piñero

... *cien años*

Se celebró un congreso internacional dedicado al estudio y proyecto de la construcción de un canal interoceánico que cruzara el istmo de América. Se eligió presidente a Ferdinand de Lesseps. En un comienzo se presentaron hasta seis vías de navegación posible, a saber: la de Nicaragua, la de Panamá, la de San Blas, la de Tiati-tolo, la de Tuyra-Caquirri-Atrato y la de Atrato-Napipi. En los primeros proyectos el canal "hipotético" de Tiati-tolo parecía el más atractivo, pero ante lo elevado de los costes presupuestados las probabilidades se fueron decantando hacia la opción del Canal de Nicaragua y el de Panamá, que fue el que en definitiva ganó. Hasta su construcción, los barcos se veían obligados a dar un rodeo por cabo de Hornos, lo que supondría un ahorro, cuando las compuertas estuvieran abiertas, de 8000 millas marinas.

...*cincuenta años*

Falleció en Barcelona Jaime Ferrán y Clúa, quizá la más polémica de las grandes figuras científicas de la España contemporánea.

Había nacido en 1852 en la localidad tarraconense de Corbera del Ebro. Hijo de un médico rural, inició su formación en Tortosa y Tarragona. Estudió después medicina en la Facultad de Barcelona, donde obtuvo el título de licenciado el año 1873. Tras una breve temporada en Pla del Penedés, se afincó en Tortosa, localidad en la que simultaneó el ejercicio de la medicina general con el de la oftalmología. Durante una década, desarrolló allí también una polifacética actividad científica y técnica, en colaboración con su amigo el químico Inocente Paulí y bajo la influencia del astrónomo y geólogo José Joaquín Landerer.

Entre sus trabajos de dicha época, se encuentran unos estudios sobre telefonía que le permitieron, en 1878, lograr con aparatos de construcción propia una comunicación telefónica entre Tortosa y Tarragona, solamente dos años después de la primera comunicación de Graham Bell. Mayor atención dedicaron entonces Ferrán y Paulí a la fotografía. En 1879, publicaron un procedimiento instantáneo basado en una emulsión de bromuro de plata y gelatina, que años más tarde motivaría un ruidoso pleito entre las firmas Young y Kodak. También realizaron investigaciones sobre medios fotográficos al carbón y a las tintas grasas, que resultaron de gran utilidad en su obra posterior.

El interés de Ferrán por la microbiología se debió inicialmente a su admiración por la obra de Pasteur, cuyos trabajos pudo estudiar en la biblioteca de Landerer, sobre todo los publicados en las *Comptes Rendus* de la Academia de Ciencias de París. El microscopio de petrógrafo que tenía el propio Landerer le sirvió para realizar sus primeros ensayos prácticos en 1880, que prosiguió después con otro que adquirió en París. Pronto se convirtió en un diestro microbiólogo y consiguió preparar las vacunas contra el carbunco y mal rojo del cerdo. En 1884, la Real Academia de Medicina de Madrid premió su *Memoria sobre el parasitismo bacteriano*. El mismo año, fue nombrado miembro de una comisión que el Ayuntamiento de Barcelona envió a Marsella, con motivo de haberse desencadenado el cólera en el sur de Francia. En colaboración con Paulí, que se desplazó a su costa, trabajó en los hospitales para coléricos de Marsella y Tolón hasta conseguir aislar y cultivar el vibrión que Robert Koch había descrito meses antes.

A su regreso a Tortosa, gracias a disponer de un cultivo virulento de vibrión colérico, Ferrán pudo provocar un cólera experimental en el cobaya y comprobar la acción inmunizadora que provocaba su inyección. Esta fue la base de su vacuna anticolérica, consistente en la inyección subcutánea de gérmenes vivos, que en principio consideró preferible a la inmunización con vibriones atenuados o muertos. Tras aplicársela a sí mismo y a una serie de voluntarios, comunicó su descubrimiento a la Academia de Ciencias de París, en marzo de 1885. Dos meses más tarde, se trasladó a Valencia, invadida entonces por la epidemia de cólera, e inició una campaña en la que vacunó a más de cincuenta mil personas. Ello motivó una dura polémica que desbordó casi inmediatamente el escenario local. Baste decir que acudieron a Valencia casi una veintena de comisiones españolas y un gran número de comisiones y personalidades científicas extranjeras, así como corresponsales de toda la prensa mundial.

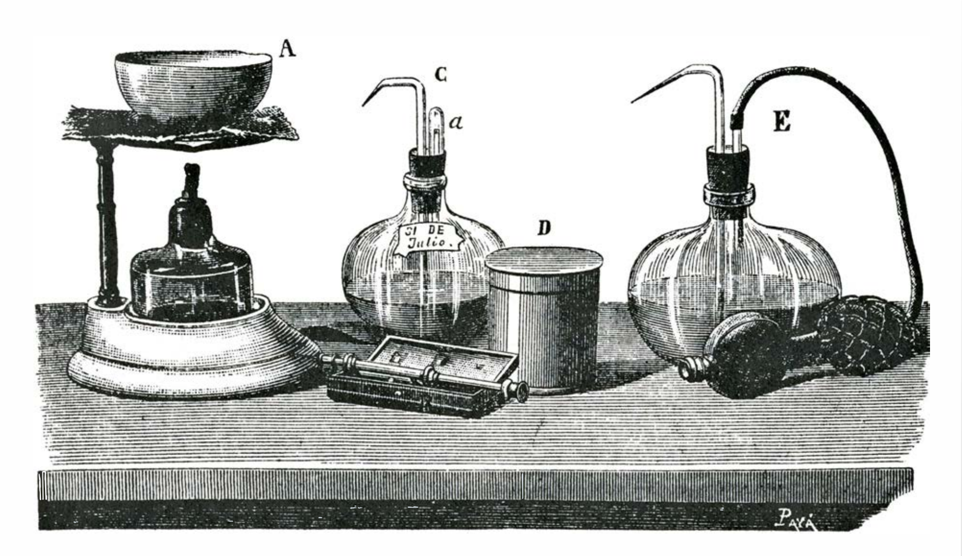
La polémica ha persistido después, en cierto modo, en los estudios históricos sobre el tema, cuyos enfoques unilaterales deforman gravemente los hechos más evidentes de la realidad histórica. No puede negarse, en primer término, la importancia de la contribución de Ferrán, que fue la primera aplicación al hombre de la vacunoterapia moderna. El reconocimiento internacional de la misma puede cifrarse en la concesión, en 1907, de la mitad del Premio Bréant, que había instituido la Academia de Ciencias de París para el descubridor de un remedio contra el cólera. Los términos del informe que, con este motivo, redactó la comisión encabezada por Pierre Roux, son muy precisos: "Es el primero que ha demostrado la acción colérica del vibrión colérico de Koch en los animales y demostrado que éstos pueden ser inmunizados... Pertenece asimismo al Sr. Ferrán la iniciativa de la inmunización preventiva del hombre contra el cólera por medio de la inyección de cultivos adecuados." Sin embargo, resultan también innegables las graves limitaciones de la aportación de Ferrán. Una audacia, cercana en ocasiones a la ligereza e incluso a la irresponsabilidad, le llevó a poner en peligro vidas humanas sin las debidas garantías y a defender, durante cierto tiempo, una morfología fantástica del vibrión o medidas preventivas pintorescas. Por otra parte, cometió el imperdonable error de intentar mantener el "secreto" de su vacuna para explotarla comercialmente.

Las dos vertientes contradictorias de la personalidad de Ferrán condicionaron igualmente su obra posterior como microbiólogo. Por iniciativa suya, el Ayuntamiento de Barcelona creó un Laboratorio Microbiológico Municipal en una fecha tan temprana como 1887. El propio Ferrán, con la aprobación explícita de Pasteur, fue el primer director del nuevo centro, al que se incorporaron como colaboradores Inocente Paulí, Luis Comenge y Ramón Turró, entre

otros. Tras numerosos incidentes, acabó destituido en 1906, a causa de sus ligerezas y abusos, que procuró exagerar la enemistad interesada de Turró, quien le sucedió en el cargo. A partir de entonces, se refugió en su laboratorio privado, al que tituló "Instituto Ferrán".

Ferrán fue el primero que preparó y aplicó la vacuna antitífica en sí mismo y en varios obreros de las alcantarillas de Barcelona (1887), pero interrumpió sus trabajos sobre el tema, sin llegar a fundamentarlos rigurosamente. Por otra parte, figuró entre los pocos bacteriólogos europeos que trabajaron en la inmunización experimental antidiftérica antes del descubrimiento del suero por Emil Adolf von Behring y Shibasaburo Kitasato (1890). También contribuyó al estudio del bacilo tetánico (1898) y al de la peste, este último con motivo de la epidemia de Oporto (1899). Se ocupó asimismo del mal rojo del cerdo, sobre todo en relación con una epizootia que hizo estragos en Mallorca (1892), e hizo notables trabajos zimotécnicos referentes a la fermentación de los vinos y de los productos lácteos. Más discutibles fueron las bases científicas de su "método supra-intensivo" de vacunación antirrábica (1888), que intentó superar al de Pasteur. Algo parecido puede decirse de sus estudios en torno a la tuberculosis, centro de su actividad a partir de 1897. Se apoyaban en una hipótesis ciclogénica del bacilo tuberculoso, consistente en mutaciones que partían de un "bacilo alfa" semejante al colibacilo, con el que pretendió preparar una "vacunación antialfa".

Ferrán publicó más de un centenar de trabajos entre libros, folletos y artículos aparecidos en revistas médicas y científicas españolas y extranjeras, sobre todo francesas y alemanas. Una de las primeras fue el folleto *La instantaneidad en fotografía* (Tortosa, 1879), que escribió en colaboración con Paulí. Su primera publicación sobre la profilaxis anticolérica fue un artículo que apareció, en 1884, en la *Gaceta Médica Catalana*. Terminada la epidemia, Ferrán, Paulí y Amalio Gimeno, catedrático de medicina en Valencia, ofrecieron una exposición sistemática sobre la vacunación en el libro titulado *La inoculación preventiva contra el cólera morbo asiático* (Valencia, 1886). Entre la producción posterior del bacteriólogo catalán hay que destacar otros dos libros: *Estudios sobre la rabia y su profilaxis* (Barcelona, 1889) y *La peste bubónica* (Barcelona, 1907), este último escrito en colaboración con F. Viñas y R. Grau. Su principal trabajo sobre el bacilo tetánico apareció en alemán, el año 1898, en el *Centralblatt für Bakteriologie* y su primera publicación acerca de la "vacunación antialfa", en francés, el año 1907, en la *Revue Internationale de la Tuberculose*.



Instrumentos utilizados por Ferrán en la vacunación anticolérica de 1885, en Valencia. **A.** *Dispositivo para esterilizar la jeringuilla;* **B.** *Jeringuilla para la inoculación;* **C.** *Matraz fechado con vacuna;* **D.** *Tarro con vacuna para aspirarla con la jeringuilla;* **E.** *Matraz dispuesto para extravasar la vacuna.*

Los efectos inmediatos y retardados de la guerra nuclear

La potencia de las fuerzas estratégicas de los Estados Unidos se fundamenta en los efectos inmediatos producidos por las armas nucleares. Igualmente importantes resultan los efectos retardados

Kevin N. Lewis

La razón de ser fundamental de las fuerzas nucleares estratégicas de los Estados Unidos está orientada a disuadir a la Unión Soviética de la idea de desencadenar un ataque contra ellos o sus aliados. Para poder cumplir esta misión, Norteamérica mantiene una constante capacidad para producir daños desastrosos en la Unión Soviética. Los misiles de largo alcance y las fuerzas de bombardeo norteamericanos han sido proyectados para poder sobrevivir incluso ante un ataque por sorpresa generalizado, en un número lo suficientemente importante como para desencadenar un devastador contraataque de represalia. Como quiera que la Unión Soviética dispone de fuerzas similares, parece improbable que ningún bando llegue a considerar ventajoso un ataque contra el otro. A ese recíproco poder de represalia, o a esa capacidad de destrucción asegurada, se suele atribuir el actual equilibrio militar estratégico entre ambas superpotencias.

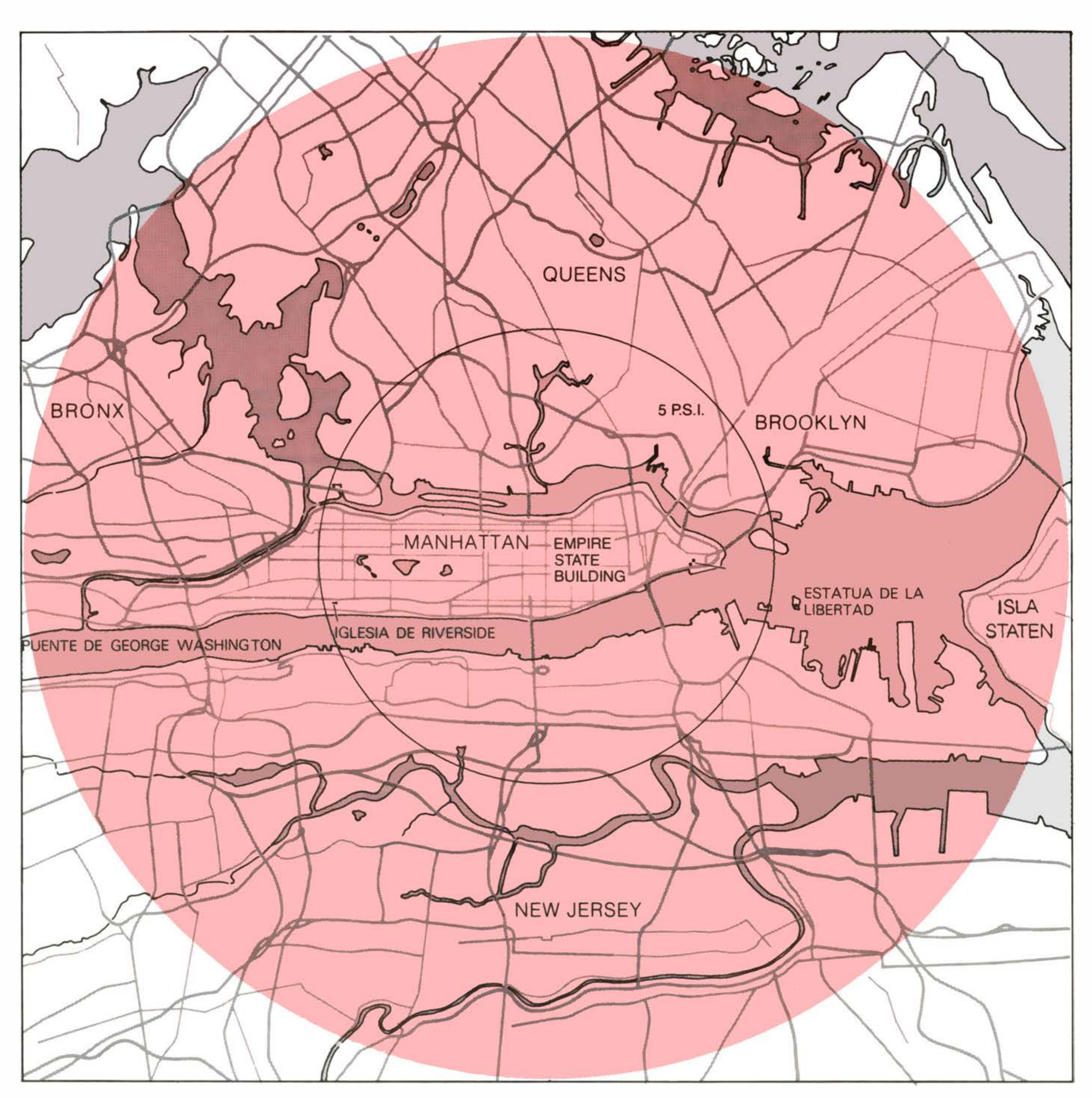
Considerando que desde esta perspectiva la evitación de una guerra depende en cierta medida de la posibilidad de mantener intacta la capacidad de destrucción asegurada de los Estados Unidos, cualquier disminución de esta capacidad constituye un asunto muy grave. De acuerdo con ello, las recientes declaraciones de algunos comentaristas militares en el sentido de que Rusia está tratando activamente de conseguir unos medios que le permitan reducir la efectividad de una acción de represalia norteamericana, han producido una preocupación generalizada. En concreto se apunta el hecho de que las actuales y ambiciosas iniciativas rusas en el terreno de la defensa civil podrían llegar a producir un peligroso desequilibrio estratégico como consecuencia de la ausencia de esfuerzos norteamericanos similares en esta materia. Por ejemplo, la Unión Soviética —conjuntamente con un ataque por sorpresa dirigido contra los misiles norteamericanos con base terrestre— podría intentar la evacuación de sus núcleos urbanos, con el resultado previsible de que, en un intercambio nuclear generalizado y devastador, las pérdidas soviéticas serían sensiblemente inferiores a las norteamericanas. En tal situación los Estados Unidos podrían llegar a quedar incapacitados para desencadenar una escalada en las hostilidades y entonces Rusia podría ser capaz, por varios procedimientos, de llegar a imponer su voluntad. Se comenta que incluso en el caso de que llegase a desencadenarse una guerra generalizada, la Unión Soviética sería capaz de recuperarse con mucha mayor rapidez que los Estados Unidos. Una de las consecuencias de este razonamiento ha sido un espectacular renacimiento del interés por el, hasta ahora agonizante, programa de defensa civil norteamericano; otro resultado ha sido la toma en consideración de nuevas opciones sobre los objetivos de los misiles estratégicos, orientándolos a atacar el programa soviético de defensa civil.

Tales consideraciones hipotéticas están basadas, en parte, en una cierta infravaloración de los daños que las fuerzas norteamericanas supervivientes podrían producir a Rusia. Muchas estimaciones de este tipo toman en consideración exclusivamente los efectos de choque, o explosivos, de las armas nucleares, calculables con facilidad. Ignoran los efectos igualmente devastadores de la radiación térmica y de la radiación de ionización. Cuando estos efectos adicionales se incluyen en los cálculos, resulta muy claro que la guerra nuclear sigue representando, sin paliativos, un desastre recíproco y que no existe hoy día nada en materia de defensa civil que pueda modificar estas perspectivas. Por todo ello, desde un punto de vista operativo militar, no resultan válidas las afirmaciones de que la capacidad de represalia norteamericana se está "erosionando". Más aún, es altamente improbable que la situación pueda llegar a cambiar en un futuro previsible.

¿Cómo se calculan los daños producidos en una guerra nuclear y qué consecuencias de una guerra de tal naturaleza suelen excluirse al hacer el cómputo de daños? En este artículo estableceremos una comparación entre los cálculos que generalmente se emplean para fijar unos niveles "adecuados" de destrucción asegurada y la estimación de las consecuencias —probablemente más graves— de una confrontación nuclear entre ambas superpotencias. El análisis más razonable pone de manifiesto que ni Estados Unidos ni la Unión Soviética tienen motivos para sufrir la más mínima preocupación acerca de su respectiva capacidad de represalia. Aunque gran parte de los actuales debates sobre la gravedad de la amenaza rusa tienden a ignorar este hecho, no cabe la menor duda de lo poco rentable que sería para ambos bandos llegar a una guerra generalizada. Por supuesto que una acción disuasoria semejante tiene su fundamento en una serie de factores en nada relacionados con la capacidad para lanzar un ataque masivo de represalia. Sin embargo, a ambos bandos interesa que el concepto de "victoria" en una guerra nuclear generalizada —en el sentido de que un bando mejore su posición relativa a unos costos aceptables— no sea tomado en consideración en el debate estratégico, y que las consecuencias totales de un acontecimiento tan calamitoso sean presentadas ante la atención pública.

Los criterios específicos sobre la eficacia de una posible represalia fueron

establecidos por vez primera bajo la dirección del entonces Secretario de Defensa, Robert S. McNamara, a comienzos de la década de los años 60. Hasta entonces, los planificadores estratégicos militares habían carecido de cualquier tipo de modelos cuantitativos formales que les pudieran servir de base para calcular unos niveles adecuados aplicables a las fuerzas norteamericanas de represalia. De esta forma, el Secretario McNamara se adelantó al concepto de destrucción asegurada manifestando que la destrucción de entre el 20 y el 25 por ciento de la población de la Unión Soviética y del 50 por ciento como mínimo de su capacidad industrial representaría, para los dirigentes de aquel país, un nivel de daños irremontable. Mediante el establecimiento de estas premisas, McNamara consiguió coordinar la planificación de las Fuerzas Aéreas y de la Armada y orientarla a conseguir el cumplimiento de las exigencias militares estratégicas sin modificar las estructuras existentes en las Fuerzas Armadas, así como también eliminar



ZONA LETAL, definida por los especialistas norteamericanos en guerra nuclear como el área circular en cuyo interior el número de supervivientes ante una explosión nuclear es equivalente al número de bajas mortales producidas en el exterior de la misma. Esta simplificación permite llegar a una estimación de las bajas mortales inmediatas que se producirían como consecuencia de una explosión nuclear, multiplicando la zona letal por la densidad de población (partiendo de la base de que la densidad de población permanece uniforme a todo lo largo de la zona geográfica). Como regla general puede considerarse que la zona letal se extiende aproximadamente dentro de un contorno sometido a una sobrepresión de hasta 34,5 kNm^{-2} (kilonewton por metro cuadrado), lo cual, para la explosión en el aire de un megatón representada en la ilustración de la página siguiente corresponde a una zona circular de unos 6,9 kilómetros de radio (zona encerrada en el *círculo negro*). El concepto de zona letal deja sin incluir algunos efectos retardados importantes en las explosiones nucleares (más difíciles de predecir), tales como los incendios consecuentes y la lluvia radiactiva. En un día despejado, una explosión de un megatón podría producir incendios a una distancia de 16 kilómetros. Si estos incendios se consolidaran hasta formar una masa de fuego, la zona dentro de su radio de acción (*zona coloreada*) resultaría devastada, multiplicando por cinco la superficie de zona letal.

1,8 SEGUNDOS DESPUES DE LA EXPLOSION

BOLA DE FUEGO

FRENTE DE LA ONDA DE CHOQUE PRINCIPAL

BRONX — PUENTE DE GEORGE WASHINGTON — IGLESIA DE RIVERSIDE — EMPIRE STATE BUILDING — WORLD TRADE CENTER — ESTATUA DE LA LIBERTAD — ISLA DE STATEN

4,6 SEGUNDOS DESPUES DE LA EXPLOSION

SOBREPRESION DE 110,3 KILONEWTON POR METRO CUADRADO

FRENTE DE LA ONDA DE CHOQUE REFLEJADA

11 SEGUNDOS DESPUES DE LA EXPLOSION

SOBREPRESION DE 41,37 KILONEWTON POR METRO CUADRADO

VELOCIDAD DEL VIENTO = 290 KILOMETROS POR HORA

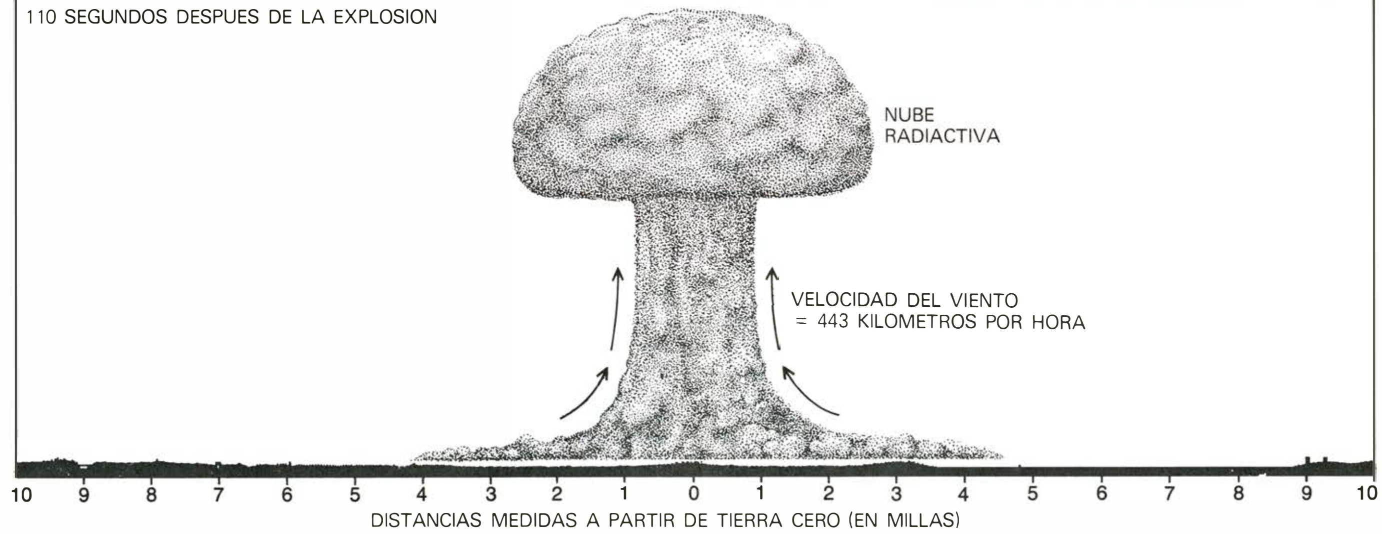
37 SEGUNDOS DESPUES DE LA EXPLOSION

RESIDUOS DE LA BOMBA EN FORMA DE GASES A ELEVADA TEMPERATURA

SOBREPRESION DE 6,9 KILONEWTON POR METRO CUADRADO

VIENTOS POSTERIORES

VELOCIDAD DEL VIENTO = 64,4 KILOMETROS POR HORA



DISTANCIAS MEDIDAS A PARTIR DE TIERRA CERO (EN MILLAS)

todo tipo de programas superfluos. Aunque la tarea de definir un determinado nivel de daños conlleva cierta implicación política, como por ejemplo amenazar al gobierno ruso con una destrucción superior a los límites de tolerancia, los porcentajes concretos elegidos reflejaban las posibilidades reales de las fuerzas estratégicas norteamericanas programadas en aquella época. También los necesarios niveles de destrucción estaban relacionados en cierta medida con las características del sistema concreto de objetivos que Rusia representaba.

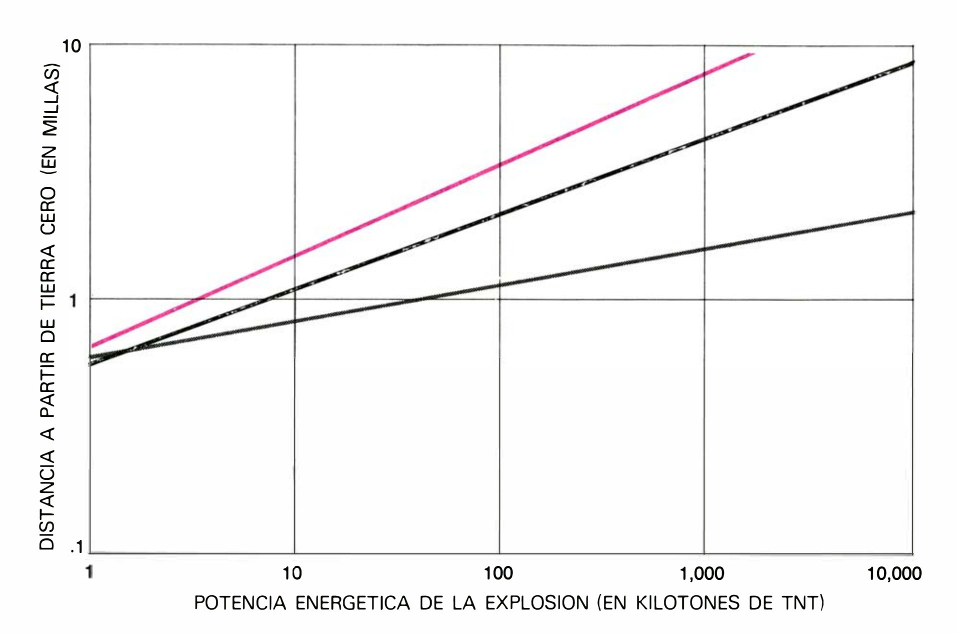
La población y los recursos económicos de la Unión Soviética están concentrados alrededor de un conjunto significativamente pequeño de grandes centros urbanos. Aproximadamente una tercera parte de la población y casi los dos tercios de la capacidad industrial están concentrados en los 200 centros urbanos más importantes del país. Un ataque nuclear dirigido contra poblaciones distintas de las anteriormente señaladas no incrementaría, de una forma apreciable, los daños de represalia (a excepción de los efectos retardados producidos por la lluvia radiactiva). Por tanto, el concepto de destrucción asegurada expuesto por McNamara podía traducirse en la capacidad necesaria para destruir las 200 ciudades soviéticas más importantes.

Partiendo de esta premisa, podrían establecerse las necesidades y exigencias de las Fuerzas norteamericanas sobre la base de determinar el número de cabezas nucleares requeridas para destruir los objetivos de importancia social y económica en las citadas 200 urbes soviéticas. El cálculo de objetivos ha de tomar en consideración una serie de factores operativos, como por ejemplo la composición y el trazado de los núcleos urbanos; pero, sobre todo, está basado en vaticinar con precisión la forma en que la población será afectada por los efectos letales de las explosiones nucleares. En la práctica real, los daños de represalia se calculan comparando las propiedades físicas de las explosiones nucleares con las características más importantes de cada una de las referidas ciudades, tomadas aisladamente como objetivo posible. En el cálculo de tales niveles de daños, los planificadores norteamericanos tienen a su disposición una enorme cantidad de información clasificada sobre cada objetivo, un conjunto muy complejo de procedimientos de análisis y una capacidad muy avanzada de proceso de datos. Sin embargo, con la ayuda de ciertos procedimientos elementales puede llegarse a conclusiones suficientemente aproximadas.

La potencia de un arma nuclear viene expresada en términos de la cantidad necesaria de explosivo químico para liberar una cantidad de energía equivalente; la potencia de un arma nuclear se mide en kilotones (millares de toneladas) o megatones (millones de toneladas) de TNT (trinitrotolueno = trilita). Al igual que sucede con una explosión química, la energía producida por una explosión nuclear se genera con mucha rapidez en un volumen muy pequeño. Cuando se produce una explosión nuclear en el aire, la energía liberada convierte en materia gaseosa los componentes de la cabeza nuclear de forma instantánea, creando una bola de fuego ardiente y de rápida expansión. La explosión produce dos clases de efectos inmediatos que pueden acarrear consecuencias devastadoras en el ataque a un núcleo urbano. En primer lugar, a medida que se produce la expansión de la bola de fuego ésta irradia una onda de choque a través del medio que la rodea. La onda de choque, que se transmite —a partir del punto en que se ha producido la explosión— a velocidades supersónicas, produce daños por el choque contra las estructuras y seres humanos. La bola de fuego en combustión genera también energía térmica, en especial fotones, en las regiones visible e infrarroja del espectro electromagnético, que pueden producir quemaduras e incendios en los materiales que no hayan sido protegidos por algún tipo de pantalla opaca. En términos aproximados, una mitad de la energía generada por el explosivo nuclear se transforma a la larga en impulsos de choque mecánico y alrededor de una tercera parte se libera en forma de radiación térmica. El resto de la energía corresponde a la radiación nuclear inmediata y a las radiaciones térmica y nuclear retardadas, ninguna de las cuales se ha considerado de importancia significativa a la hora de estimar la destrucción asegurada, aunque sin duda cooperan realmente a incrementar el poder destructor de un ataque nuclear.

Los impulsos mecánicos de una explosión nuclear son similares a los producidos por las mareas. El frente de choque opera como un muro de aire comprimido. A su paso, todo tipo de estructuras queda expuesto a una elevación casi instantánea de la presión atmosférica local y pueden ser aplastadas. A continuación del frente de choque, se producen fuertes vendavales, análogos a las corrientes de agua que siguen a una ola oceánica en movimiento. La energía producida por estos huracanes puede causar igualmente el derrumbamiento de las estructuras en la zona del objetivo. Según sea la configuración y el tipo de las construcciones, los edificios pueden resultar vulnerables a la onda de choque, a los vientos que vienen a continuación o

LOS EFECTOS INMEDIATOS de la explosión de una cabeza nuclear de un megatón, producida a una altura de 1981 metros por encima del corazón de Nueva York, aparecen descritos cronológicamente en la serie de figuras de la página anterior. Inmediatamente después de la explosión se formará una bola de fuego extremadamente caliente y luminosa. Esta bola de fuego emitirá una radiación térmica muy intensa (*color*), capaz de producir quemaduras en la piel y de iniciar incendios a considerable distancia. La explosión dará lugar igualmente a una onda de choque arrasadora que se transmitirá a partir del centro de la bola de fuego a velocidad supersónica; por ejemplo, a los 1,8 segundos de haberse producido la explosión, el perímetro de la onda de choque (*círculo negro*) se encontrará aproximadamente a 0,805 kilómetros del centro de la bola de fuego. Además, los procesos nucleares consecuentes a la explosión serán acompañados por la emisión de radiación, fundamentalmente rayos gamma y neutrones (*líneas blancas ondulantes*), de un alcance, en el aire, suficiente para llegar hasta la superficie, alcanzando los límites de la zona media de la ciudad. Cuando la onda de choque principal llegue a la superficie terrestre se producirá, por reflexión, una segunda onda de choque. A cierta distancia de "tierra cero" (dependiendo de la altura de la explosión y de la potencia del arma) los frentes de ambas ondas de choque —principal y reflejada— se funden en uno solo, cerca de la superficie, para constituir un único frente reforzado Mach; en el caso de una cabeza nuclear de un megatón detonada a 1981 metros, el efecto Mach comenzará alrededor de 4,6 segundos después de la explosión, a una distancia de 1,3 millas (2,0921 kilómetros) de tierra cero. En dicho punto la sobrepresión (es decir, la presión del aire por encima de la presión atmosférica ambiental) será de 110,316 kilonewton por metro cuadrado (kNm^{-2}). A los 11 segundos de la explosión, el frente Mach se habrá trasladado a 3,2 millas (5,15 kilómetros) de tierra cero, la sobrepresión en el frente Mach será de 41,369 kNm^{-2} y la velocidad del viento justamente siguiendo al frente será de unos 289,682 kilómetros por hora; una considerable cantidad de radiación térmica y nuclear continuará llegando hasta la superficie. A los 37 segundos de la explosión el frente Mach se encontrará aproximadamente a 9,5 millas (15,12 kilómetros) de tierra cero, la sobrepresión a la altura del frente será de 6,895 kNm^{-2} y la velocidad del viento detrás del frente será de unos 64,35 kilómetros por hora (el cristal se rompe a sobrepresiones hasta de 34,47 kNm^{-2}). Aunque la radiación térmica habrá dejado de ser significativa, los rayos gamma todavía alcanzarán la superficie en cantidades potencialmente letales. La bola de fuego ya no será luminosa, pero todavía se mantendrá a temperaturas muy elevadas y por ello ganará altura con mucha rapidez, produciendo fuertes corrientes de aire hacia su interior y de abajo arriba. Estos vientos arrastrarán en su movimiento ascendente y centrípeto toda clase de polvo, basura y escombros de la ciudad, formando el tronco de lo que pasará a constituir a la larga la característica nube en forma de hongo. Alrededor de 110 segundos después de la explosión, los residuos calientes de la bola de fuego, a la vez que siguen elevándose, habrán comenzado a expandirse y a enfriarse. Como consecuencia, los productos de la fisión, en forma gaseosa, así como otros residuos del explosivo, se condensan para formar una nube de partículas radiactivas. En estos momentos la nube se habrá elevado a una altura de 11,265 kilómetros. La máxima altura que alcanzará la nube (después de diez minutos) será de aproximadamente 22,531 kilómetros. Por último, las partículas componentes de la nube serán dispersadas por el viento y, a menos que se produzcan precipitaciones atmosféricas, no caerá una temprana (o local) lluvia radiactiva.



RADIOS DE ACCION TIPICOS de tres tipos diferentes de efectos dañinos consecuentes a las explosiones nucleares. Corresponden a una típica explosión en el aire, y se dan en función de la potencia energética de dicha explosión. La recta de color indica la distancia tope hasta la que la radiación térmica puede producir quemaduras de segundo grado e iniciar incendios con riesgo de producir un fuego masivo. La recta negra mide el radio del círculo correspondiente a la sobrepresión de los 34,5 kNm^{-2}, dentro del cual el paso del frente de la onda explosiva seguida de un viento de hasta 257,5 kilómetros por hora podría producir destrucciones urbanas masivas y un elevado porcentaje de bajas mortales. La recta de color gris indica el radio de acción dentro del cual se producirá un 100 por ciento de bajas mortales como consecuencia de la radiación nuclear inmediata ocasionada por la explosión. Resulta evidente que en condiciones atmosféricas favorables los efectos destructivos de carácter térmico correspondientes a esta explosión podrían alcanzar bastante más allá de la zona de daños importantes producidos por la onda de choque. La radiación nuclear inmediata, por otra parte, no resulta ser, de ningún modo, una fuente productora de daños de consideración en términos de explosivos nucleares estratégicos (los cuales pueden oscilar entre varias decenas de kilotones hasta varios megatones), ya que las áreas cubiertas por la radiación mortal se encontrarán igualmente expuestas a los efectos de choque y térmicos. Unicamente en los casos de potencias explosivas mucho menores (del orden de un kilotón o aún menores) la radiación nuclear inmediata se convierte en origen importante de bajas mortales; esta circunstancia, en realidad, constituye el principio básico de las armas de radiación perfeccionada, o de la bomba de neutrones. (*Investigación y Ciencia*, julio 1978.)

a ambos a la vez. La "dureza" de un objetivo (su capacidad de resistencia ante los efectos destructores de la onda de choque) suele describirse en términos de la máxima "sobrepresión" inducida (en libras por pulgada al cuadrado por encima de la presión atmosférica –una libra por pulgada al cuadrado equivale a 6,8950 kilonewton por metro cuadrado–) necesaria para lograr la destrucción del objetivo.

La radiación térmica puede producir bajas por quemaduras, bien directamente por el destello de la explosión, o bien indirectamente (a través de la combustión de los materiales circundantes) por incendio, superponiéndose ambos efectos a las bajas producidas por el efecto de choque. La severidad de tales daños depende en ambos casos de la potencia de la energía radiante liberada (medida en general en calorías por centímetro cuadrado) y del período de tiempo en que se ha liberado esta energía. Los efectos destructivos del choque disminuyen con la distancia mucho más rápidamente que los efectos térmicos. Por tanto, en condiciones óptimas, una explosión nuclear puede producir daños por incendio, de gran importancia, mucho más allá de la zona asolada por el efecto de choque. Sin embargo, los daños térmicos se verán condicionados en gran manera por factores externos, entre los que se incluyen la presencia de nubes o de nieve, la transparencia relativa de la atmósfera y las características del objetivo. Por ello, los efectos térmicos son mucho más difíciles de predecir con exactitud que los efectos de choque.

Puesto que las fuerzas de represalia se calculan sobre la base de daños asegurados, la consecuencia de un ataque se acostumbra a calcular sólo en razón de los efectos de choque, es decir, de aquellos que pueden predecirse con mayor exactitud. Consideremos el problema de la asignación de un "paquete" adecuado de armas nucleares a emplear contra una determinada zona urbana, después de haber pasado revista a los diferentes subobjetivos que constituyen el objetivo total representado por dicha ciudad. Las distintas zonas de impactos para las diferentes armas se seleccionarán de forma tal que permitan asegurar que la integridad del objetivo quede barrida por el efecto de choque total. Si los subobjetivos están lo bastante cercanos unos de otros, puede bastar con una cabeza nuclear. Si los subobjetivos están dispersos o protegidos, resulta más conveniente la asignación de más de un arma aislada a una zona que, por contra, incrementar la potencia de un proyectil único. Este procedimiento evita el peligro del posible fallo en la explosión de una cabeza nuclear aunque sea de gran potencia, lo que comportaría dejar un subobjetivo sin batir. También está en relación con el hecho de que pocos de los complejos industriales o militares se encuentran lo suficientemente concentrados o tienen la configuración adecuada, para ser atacados por un proyectil aislado del tipo habitual que hay en los arsenales estratégicos norteamericanos.

Cada ciudad posee un conjunto muy diferenciado de características propias y específicas como posible objetivo. Sin embargo, existen algunas reglas muy sencillas que permiten predecir los daños y las bajas. En términos generales, cualquier tipo de estructura que no haya sido fabricada exprofeso para resistir la onda de choque será destruida si se le somete a una sobrepresión de unas 5 o más libras por pulgada al cuadrado (p. s. i., unos 34,5 kilonewton por metro cuadrado) por encima de la presión atmosférica ambiental de unos 15 p.s.i., (unos 96,53 kilonewton por metro cuadrado), y aquellas estructuras que no hayan resultado destruidas realmente quedarán dañadas muy por encima de cualquier posible reparación. Algunos edificios muy reforzados (y los equipos pesados instalados en su interior) podrían resistir una sobrepresión de 40 p. s. i. (unos 275,8 kilonewton por metro cuadrado) o incluso mayor; pero si estos objetivos fuesen considerados importantes, un presunto agresor podría disminuir la altura prevista para que sus armas hicieran explosión, o podría dirigir sus armas (o asignar nuevas cabezas nucleares) para conseguir los efectos de destrucción deseados. En general, y como regla aproximada, puede afirmarse que una sobrepresión de 5 p. s. i. resulta suficiente para lograr la destrucción de la mayoría de las estructuras.

El cuerpo humano puede soportar un choque de mucha mayor intensidad que la generalidad de los edificios. Por tanto, en un ataque nuclear la mayor parte de las bajas producidas por los efectos de choque lo serán como consecuencia de efectos indirectos. El riesgo para la mayoría de la población se producirá como consecuencia de estar en el interior o en las cercanías de edificios que se derrumban, al ser alcanzados por los escombros lanzados a su alrededor por la

onda de choque, o al ser arrastrados y golpeados contra superficies inmóviles. Los efectos térmicos producirán igualmente un gran número de bajas dentro de un determinado radio de acción, con independencia de las condiciones externas. En el cálculo de bajas se emplea con frecuencia el concepto simplificado de "zona letal". Por "zona letal" –basándose en los datos teóricos y empíricos desarrollados por la Comisión de Energía Atómica en 1950– se entiende la zona circular en cuyo interior el número de supervivientes resultaría igual al número de bajas producidas en su parte exterior, considerando que la densidad de población permanece uniforme.

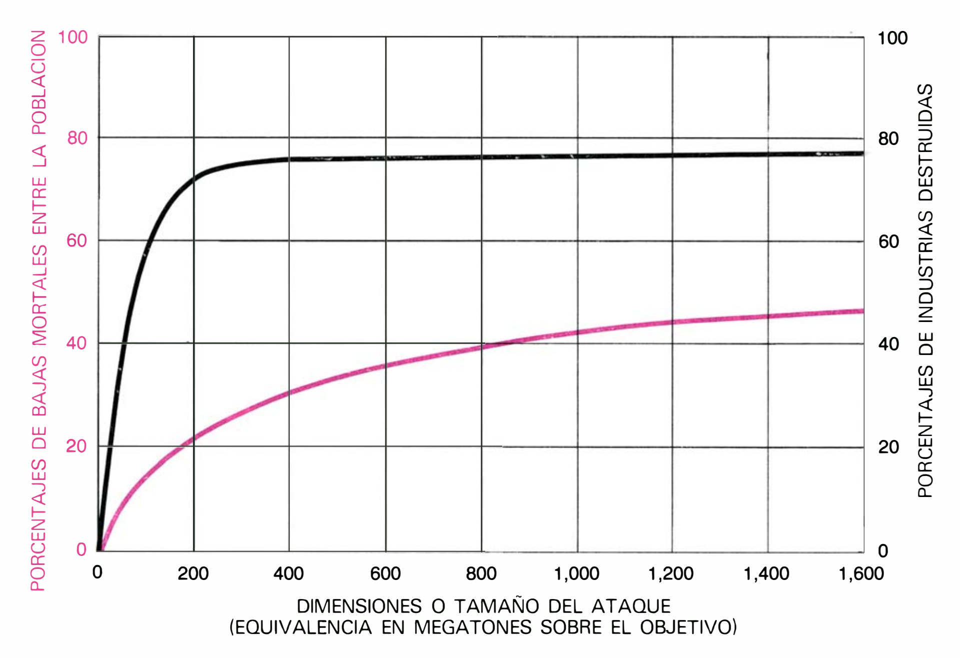
Para simplificar los cálculos, las bajas mortales se vuelven a distribuir de acuerdo con el siguiente criterio: los planificadores consideran que todas las bajas producidas en el interior de dicha zona circular son mortales y ninguna de las producidas en su exterior lo es. De acuerdo con ello, se hace un cálculo de las bajas inmediatas multiplicando la zona letal por la densidad de población. La experiencia de Hiroshima y Nagasaki, junto con los datos extraídos de las diferentes pruebas realizadas, indican que, para unas armas de alrededor de los 20 kilotones, la zona letal se extiende aproximadamente hasta unos límites que vienen a coincidir con el contorno dentro del cual se produce una sobrepresión de 34,5 kilonewton por metro cuadrado (kN m^{-2}.) Así, la superficie cubierta por dicha sobrepresión se considera una medida satisfactoria para calcular no sólo las bajas mortales, sino también el poder destructivo de las explosiones nucleares que puedan llegar a ser consideradas rentables.

Las armas nucleares generan una sobrepresión de 34,5 kN m^{-2} a una distancia que es proporcional a la raíz cúbica de su potencia. Por esta razón se afirma que cuanto mayor es el arma, con menor efectividad distribuye su poder destructor. Por ejemplo, una bomba de 100 kilotones generará una sobrepresión de 34,5 kN m^{-2} a una distancia de unos tres kilómetros. Pues bien, una cabeza nuclear de una potencia explosiva diez veces mayor (un megatón) generará una sobrepresión idéntica, pero sólo a una distancia dos veces superior. Reconociendo pues la mayor eficacia, en proporción, de las armas de menor potencia se ha llegado a un sistema gradual de medidas conocido con el nombre de "equivalencia en megatones" –definido como la potencia de una bomba en megatones elevado a los dos tercios–. Este se considera un índice más adecuado para medir la capacidad de ataque contra ciudades, que no una posible elección, muy poco precisa por lo general, de megatones. El equipo de análisis de sistemas de McNamara calculó en 1960 que el lanzamiento de una equivalencia de 400 megatones podría acabar con el 30 por ciento de la población de la Unión Soviética y destruir el 75 por ciento de su capacidad industrial. Más recientemente, los daños previsibles sobre la población y la industria han sido reajustados.

En realidad estos niveles de daños constituyen los mínimos resultantes de una explosión nuclear, ya que han sido calculados sólo sobre la base de los efectos inmediatos previsibles, de los que nos hemos ocupado antes. Al tomar en consideración los efectos retardados (incendios, lluvia radiactiva, etc.), la estimación de daños resulta mucho más elevada. Los efectos retardados aseguran también que, aun en el caso de no poder llegar a alcanzarse los niveles de daños por choque que se citan en las definiciones de destrucción asegurada, una guerra nuclear generalizada podría acabar desembocando en la destrucción completa de ambos países contendientes.

Los efectos inmediatos y retardados producidos por las armas nucleares pueden someterse a verificación si consideramos un ataque contra un objetivo urbano típico, por ejemplo, la zona metropolitana del gran Boston. La detonación de diez cabezas nucleares de un megatón cada una, dirigidas contra unos blancos elegidos entre determinados centros locales económicos y militares, generaría una sobrepresión de 34,5 kN m^{-2} en una zona de una superficie superior a los 805 kilómetros cuadrados. Resultarían muertos más de 1,3 millones de seres humanos a consecuencia de los efectos inmediatos de choque y térmicos de las explosiones y quedaría inutilizada más del 80 por ciento de la capacidad industrial de la zona. Es muy probable que los efectos secundarios de las explosiones, en especial los producidos por los fuegos y por la lluvia radiactiva, incrementasen notablemente estos porcentajes.

Si las condiciones para el ataque nuclear fuesen favorables, los efectos más devastadores serían, a buen seguro, los de carácter incendiario. En determinadas condiciones atmosféricas, cada explosión de un megatón podría iniciar fuegos hasta una distancia de unos 16 kilómetros. En un ataque como el del ejemplo anterior, la amenaza de daños producidos por el fuego abarcaría gran parte de la región este de Massachusetts. Los fuegos inducidos se sumarían a los producidos por el derrumbamiento y estallido de calderas, hornos y estufas. Los



CRITERIO DE DESTRUCCION ASEGURADA en que se apoyan los planificadores estratégicos norteamericanos para determinar el potencial de represalia requerido por las fuerzas nucleares de los Estados Unidos para disuadir y prevenir un ataque por sorpresa soviético. Dicho criterio aparece aquí expresado en función del número de "megatones equivalentes" que sería necesario lanzar para destruir los centros clave de población y los objetivos industriales de la Unión Soviética. (Equivalencia en megatones se define como la potencia explosiva de un arma nuclear elevada a los dos tercios de su potencia.) Dado el valor decreciente del incremento adicional de la equivalencia en megatones para un ataque de represalia de naturaleza tal, resulta evidente a la vista de estas curvas que el lanzamiento sobre el objetivo de unos 400 megatones equivalentes sería más que adecuado para conseguir los efectos de destrucción asegurada. Los daños en la población (*curva de color*) fueron estimados tomando en consideración los efectos mortales; los daños industriales (*curva negra*) fueron determinados calculando el "valor añadido manufacturado" destruido durante un ataque de represalia. (Valor añadido manufacturado es el incremento de valor conferido a los productos en bruto en cualquier proceso industrial.)

escombros incandescentes esparcidos y la ruptura de cisternas, depósitos y oleoductos añadirían combustible a los incendios. Las zonas de cortafuegos serían salvadas por los materiales arrojados por la explosión. Después del ataque, el intento de extinguir cientos y cientos de incendios aislados y de escasa entidad podría llegar a resultar un trabajo monumental; los puntos de toma de agua habrían saltado hechos pedazos y los equipos de lucha contra incendios habrían quedado destruidos o desarbolados. En Hiroshima, alrededor del 70 por ciento de los equipos urbanos de lucha contra incendios quedaron aplastados como consecuencia del derrumbamiento de los cuarteles de bomberos, y el 80 por ciento de los bomberos no pudieron presentarse en sus puestos.

Según sean las condiciones atmosféricas y las características de la zona del objetivo (en especial la densidad de las estructuras inflamables), los numerosos puntos de iniciación de fuegos, al principio aislados, pueden llegar a consolidarse en dos tipos diferentes de fuegos masivos: una tormenta de fuego y una conflagración. La primera aparece como consecuencia de una succión muy intensa de aire caliente, de abajo arriba, reemplazada por una corriente de aire frío arrastrado desde la periferia de la zona del fuego. Una conflagración puede surgir como complemento de la precedente, a causa de un fuerte viento terrestre que hubiese hecho aparición con anterioridad al ataque nuclear. Mientras una tormenta de fuego mantiene su actividad tan sólo durante la acción centrípeta de las corrientes señaladas, una conflagración dura mientras disponga de materiales combustibles.

La consecuencia de los fuegos masivos es la devastación más absoluta a todo lo ancho de la zona afectada. Las temperaturas producidas por un fuego masivo pueden exceder los 1000 grados Celsius, es decir, una temperatura superior a la necesaria para fundir vidrios y metales y para producir la combustión de materiales a prueba de incendios. En Hiroshima, una bomba atómica del orden de los 15 kilotones ocasionó una tormenta o explosión de fuego que permaneció en actividad durante seis horas, destruyendo por completo una parte de la ciudad de una superficie de 7,08 kilómetros cuadrados. Las urbes americanas están construidas con materiales más resistentes al fuego que las de Hiroshima; sin embargo, también es cierto que las ciudades americanas tienen más altura en sus edificios y hay repartidos numerosos puntos de distribución de carburantes —en especial gasolinas y aceites— ideales para alimentar los fuegos que puedan haberse provocado. Otro factor preponderante consiste en el hecho de que las modernas armas nucleares estratégicas exceden en dos o más órdenes de magnitud a la potencia de las que hicieron explosión en Hiroshima y Nagasaki. Además, gran parte de la zona sometida al ataque quedaría bajo los efectos de la radiación térmica producida por más de una bola de fuego.

Los refugios contra la onda de choque proporcionan escasa protección contra fuegos de consideración. La supervivencia de los ocupantes de tales refugios dependerá fundamentalmente de la temperatura y la humedad en el interior del lugar. Si llegara a desencadenarse un fuego masivo, el problema de conseguir mantener en el interior del refugio condiciones del medio ambiente en las que el personal ocupante lograse sobrevivir, se vería agravado hasta unos límites que quizá desbordasen toda posible solución. Además, a no ser que se dispusiera de un sistema de suministro de oxígeno in-

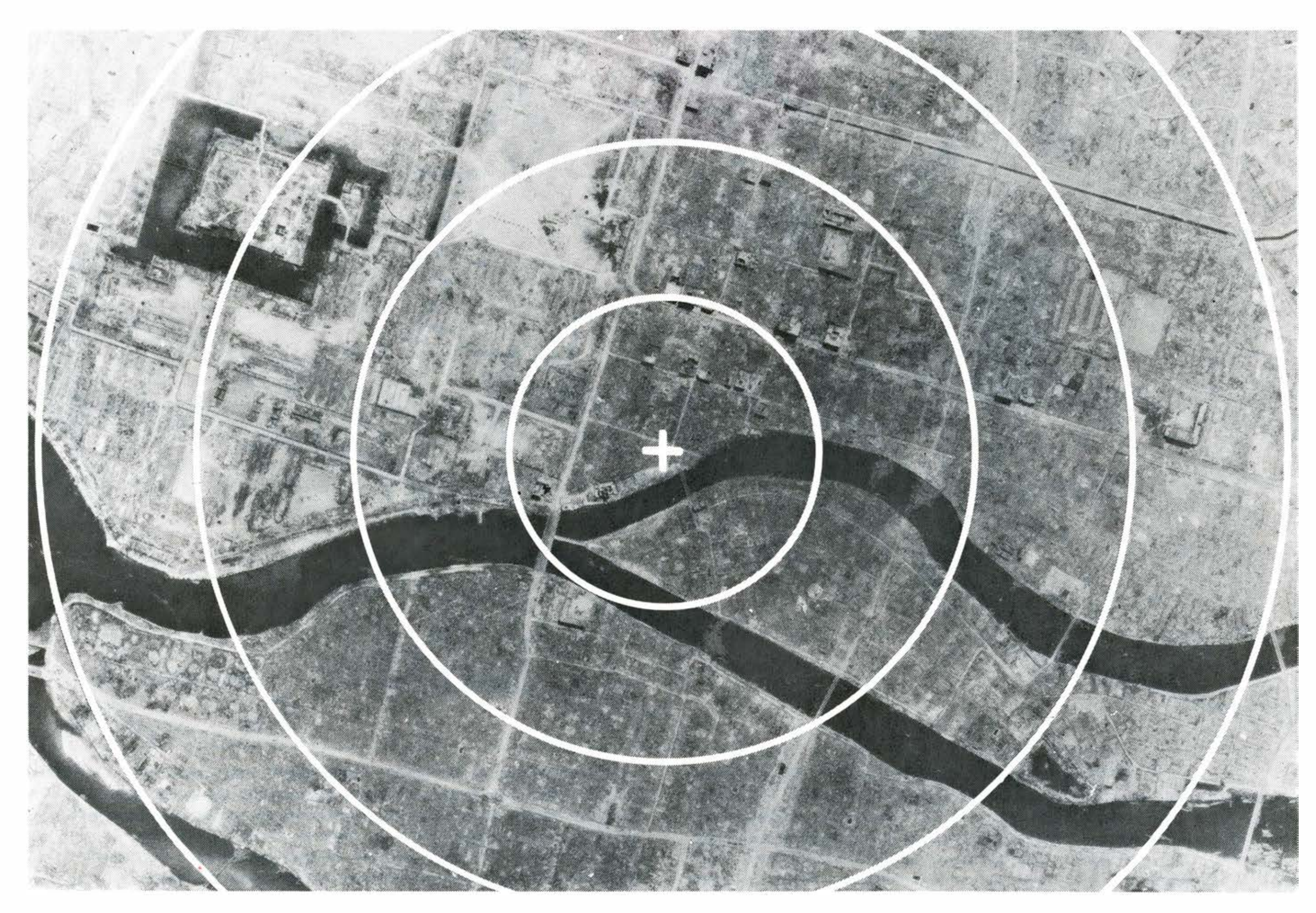


HIROSHIMA se presenta en esta perspectiva vertical formada por fotografías de reconocimiento, antes (*izquierda*) y después (*derecha*) del bombardeo atómico el 6 de agosto de 1945. La cruz señala la Tierra Cero, es decir, el punto del terreno directamente en la vertical de la explosión. Los

dependiente para cada refugio, el monóxido de carbono y otros gases tóxicos generados por el fuego podrían resultar mortales para sus ocupantes. El recalentamiento de estos abrigos, tanto por las llamas como por los cascotes y escombros incandescentes (que pueden mantenerse con una temperatura irresistible durante varios días después de la extinción del fuego) pueden exponer a los ocupantes de estos abrigos a unas condiciones ambientales en que resulte imposible la supervivencia. En Dresde, donde una tormenta de fuego producida por bombas con carga química produjo la muerte de más de 100.000 seres humanos en 1945, sólo aquellos individuos que habían abandonado sus abrigos antes de que el fenómeno se iniciase lograron sobrevivir a las dos amenazas que les acechaban: los gases tóxicos y el recalentamiento de los refugios.

Después de un ataque nuclear, un gran número de seres humanos quedarán incapacitados, atrapados bajo los edificios que se derrumben e imposibilitados para huir de la ciudad a causa de los obstáculos alzados en las calles por los escombros y el fuego. Si se llegasen a formar fuegos masivos, lo cual es una consecuencia probable de un ataque realizado a base de diversas explosiones del tipo del megatón o superiores, los supervivientes entre los que hayan logrado zafarse de la incapacitación inmediata pueden ser muy escasos. Si en el ejemplo de Boston, visto anteriormente, llegasen a consolidarse fuegos masivos, el número de bajas mortales se podría incrementar en unos 500.000 seres humanos.

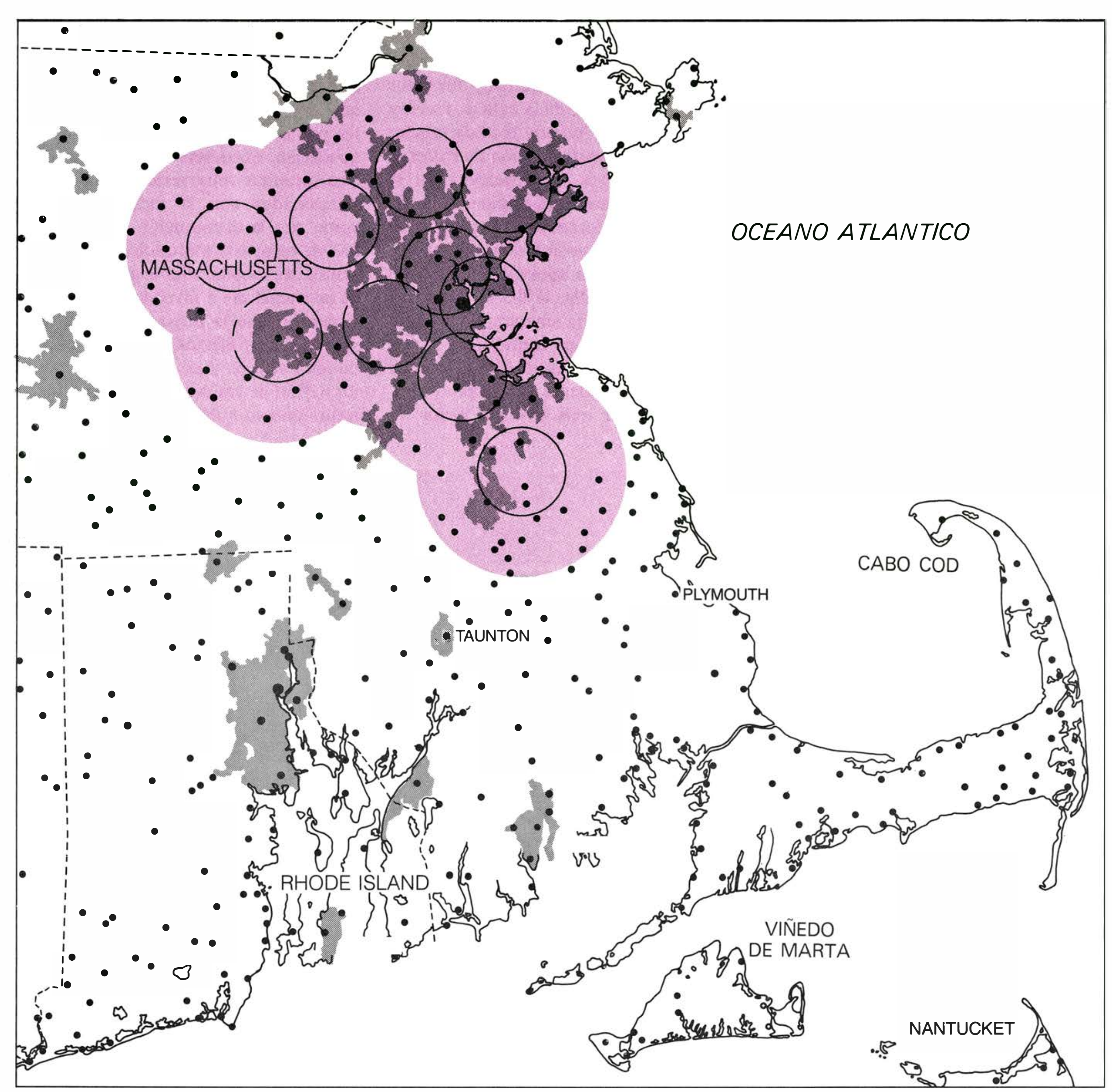
Otro factor no incluido en muchos cálculos de destrucción asegurada es el relativo a la lluvia radiactiva. Esta procede de la condensación de los subproductos radiactivos de una explosión nuclear sobre los materiales fundidos por el intenso calor de la bola de fuego y también (aunque en una medida mucho menor) de la transformación de materiales no radiactivos en radiactivos por la absorción de neutrones originados en las radiaciones nucleares de la explosión. Si se produce la explosión de un arma nuclear a nivel de la superficie terrestre o muy cercano a ella, la lluvia radiactiva se convierte en una amenaza de primera magnitud. Grandes cantidades de escombros serán sumergidos en el interior de la nube ascendente que se producirá, para descender más tarde (o mejor dicho, para ser filtrados) a través de dicha nube y desparramarse en cantidades letales a lo largo de cientos de kilómetros siguiendo la dirección del viento. Una dosis de radiación ionizante que oscile entre los 400 y los 500 rems (índice para medir los efectos biológicos de los diferentes tipos de radiación sobre el hombre), y actuando a lo largo de un período de varios días, produciría la muerte de la mitad del personal sometido a su exposición. Una dosis entre los 200 y 300 rems causaría efectos mortales a algo menos del 20 por ciento (dando por supuesto un tratamiento médico inmediato), aunque habría que considerar que podría producirse un incremento adicional en el número de bajas mortales por las graves infecciones a las que quedarían expuestas muchas personas con síntomas muy serios en la sangre, relacionados con la radiactividad, así como con la disminución en su capacidad de respuesta inmunológica. Si se hicieran explosionar diez armas –cada una de un megatón– a nivel del suelo (condiciones en que la lluvia radiactiva tiene sus máximos efectos, muy por encima de los térmicos y de choque), al menos un millón de residentes de Nueva Inglaterra –que no hubieran quedado expues-



círculos concéntricos están separados entre sí 304,8 metros. La explosión o tormenta de fuego, consecutiva a los efectos inmediatos de la explosión nuclear, permaneció en actividad durante unas seis horas. La bomba destruyó en su totalidad una superficie de la ciudad de 7,08 kilómetros cuadrados.

tos a los efectos inmediatos de choque y térmicos de las explosiones– sufrirían ahora niveles peligrosos de radiación. Incluso haciendo consideraciones optimistas sobre la eficacia y disponibilidad de refugios y de provisiones alimenticias, las bajas mortales por lluvia radiactiva, que habría que añadir a las calculadas para la zona de Boston, podrían alcanzar la cifra de los 500.000. Un ataque de este tipo dirigido contra núcleos urbanos podría combinar las explosiones en el aire con las explosiones a ras de tierra.

El número de bajas mortales producidas por el fuego y la radiación se irá incrementando de una forma constante después de un ataque de tal naturaleza, en parte porque los medios y el personal sanitario quedarán mermados. Los casos de quemaduras pueden plantear un problema médico de excepcional importancia, ya que los casos graves de quemados requieren un tratamiento intenso e inmediato para conseguir su supervivencia. La capacidad de cualquier sistema de asistencia sanitaria para el manejo de grandes cantidades de este tipo de bajas, incluso en tiempo de paz, resulta siempre considerablemente limitada. La llegada de alrededor de 50 supervivientes del accidente aéreo sufrido hace un par de años por dos grandes aviones de línea regular en Tenerife (Canarias) puso en tensión a los centros de quemados de los Estados Unidos, con una capacidad

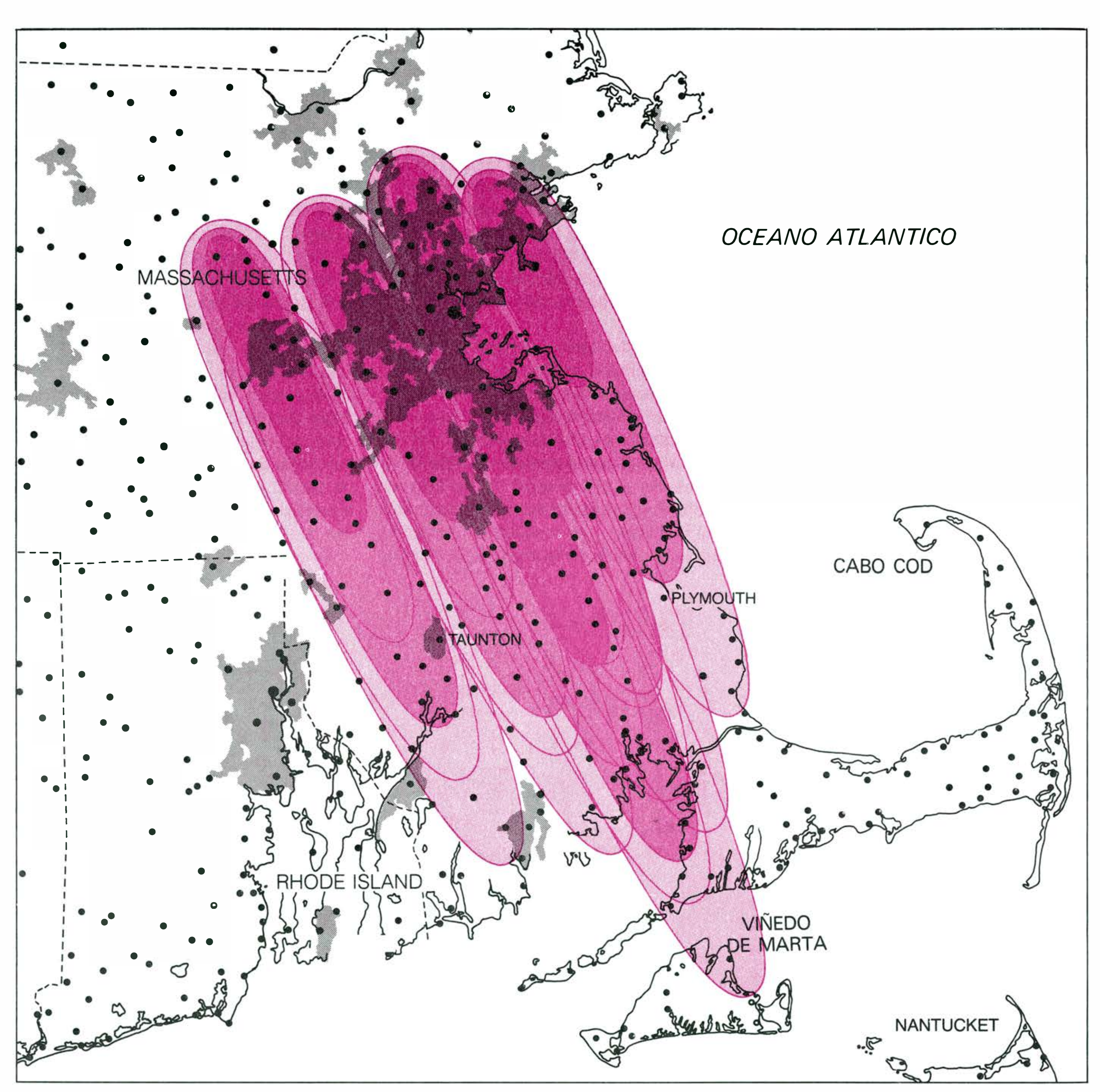


UN ATAQUE IMAGINARIO sobre la zona metropolitana del gran Boston, representado en estas dos páginas, sirve para comparar los efectos inmediatos y los retardados de una serie de explosiones nucleares. En ambos casos representados, el ataque consiste en la explosión de diez cabezas nucleares de un megatón, dirigidas contra objetivos económicos y militares. En la ilustración de la izquierda se ha supuesto que todas las armas han sido hechas detonar a una altura seleccionada como la óptima para conseguir un máximo de efectos de choque y térmicos. Las zonas encerradas en círculos de color negro corresponden a zonas expuestas a una sobrepresión de al menos 34,5 kNm^{-2}; cada una de estas zonas tiene 6,9 kilómetros de radio. Las zonas coloreadas representan las zonas expuestas a un riesgo de fuego importante y de quemaduras en un día despejado; cada zona de estas características tiene un radio de 16,093 kilómetros. El principal efecto retardado en un ataque como el aquí representado es el riesgo de que aparezca una explosión o tormenta de fuego o una conflagración de grandes dimensiones, lo que podría dar lugar a unas 500.000 bajas mortales a añadir a las previstas en el cálculo de destrucción asegurada –estimadas en 1,3 millones de muertos–, a causa de los efectos inmediatos de choque y

máxima cada uno de alrededor de los 130 pacientes. Después de un ataque nuclear, el número de bajas por quemaduras será mayor.

Los elementos sanitarios disponibles se verán muy sobrecargados de trabajo por la coincidencia de muy diversos tipos de heridas de muy distinto origen y gravedad. El peligro de heridas ocasionadas por objetos lanzados por el aire (en especial fragmentos de vidrio de ventanas hechas añicos) puede presentarse a más de 13 kilómetros del punto de explosión de un arma de un megatón y serán frecuentes las quemaduras graves hasta una distancia de 14,5 kilómetros, según las condiciones atmosféricas. Muchas víctimas de quemaduras, de enfermedades producidas por la radiactividad, o de otras causas mortales, que no mueran de inmediato, requerirán unas atenciones médicas intensivas (que en estas circunstancias resultará imposible proporcionarles). La atención al personal con lesiones de menor gravedad, a los niños, a los ancianos y a aquellos otros que requieren cuidados médicos especiales se verá complicada.

Entre los supervivientes a un ataque nuclear generalizado habrá muchos que hayan quedado con una incapacidad permanente por mutilación de miembros, cegueras u otras causas. Cualquier esfuerzo médico –por grande que éste sea– resultará degradado y falto de efica-



térmicos de la explosión. En la ilustración de la derecha se ha supuesto que la totalidad de las diez cabezas nucleares han sido hechas explosionar a nivel de la superficie del terreno, con objeto de lograr los máximos niveles de efectos por lluvia radiactiva (se han supuesto unos vientos típicos de enero al dibujar los contornos). Las zonas de color intenso son las cubiertas por una cantidad de radiación fatal para un 80 por ciento al menos de la población expuesta. Las zonas de un tono de color medio son aquellas en las que por lo menos un 50 por ciento de la población al descubierto moriría o sufriría lesiones y enfermedades relacionadas con la radiactividad. Las zonas de color suave representan los límites probables de la zona dentro de la cual aparecerían señales evidentes y síntomas clínicos de radiación en gran parte de la población expuesta, produciendo la muerte de quizás un 20 por ciento. (Es probable que los supervivientes sufriesen los efectos de la lluvia radiactiva aparecida posteriormente a consecuencia de los ataques producidos en regiones adyacentes.) El número total de bajas en la zona de Boston, consiguientes a un ataque en el que se hubiesen combinado las explosiones en el aire y en la superficie, podrían exceder de los dos millones de muertos, con un número afín de bajas humanas por heridos o enfermos.

cia por la destrucción del personal y de la infraestructura previstos para el mantenimiento de la salud pública, la aparición y proliferación de organismos causantes de enfermedades diversas (que tienden además a sobrevivir incluso ante altos niveles radiactivos) y muchas otras dificultades, como puede ser por ejemplo el aparentemente insoluble problema del manejo y medidas a adoptar con los fallecidos. El total de bajas que podría producirse como consecuencia de un ataque contra la región de Boston realizado con armas nucleares, tanto de explosión en el aire como a ras de tierra, podría exceder ampliamente los dos millones de muertos, con aproximadamente el mismo número de heridos o enfermos.

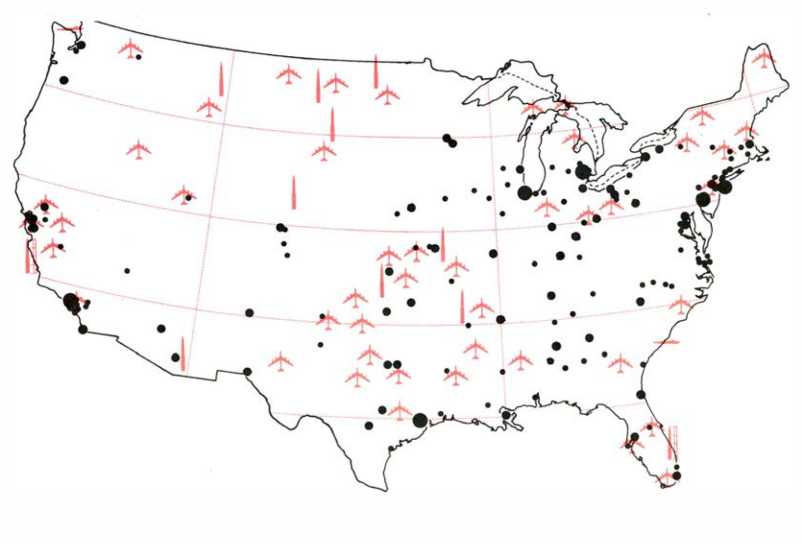
El concepto de destrucción asegurada tampoco toma en consideración algunos aspectos estratégicos. Se comenta a veces que la Unión Soviética está tratando de llevar a la práctica dos tipos de programas encaminados a conseguir que la efectividad de un posible ataque norteamericano de represalia caiga en el vacío. El primer programa está basado en conseguir la reducción del número de cabezas nucleares norteamericanas antes de que éstas alcancen sus objetivos; para ello tratarán de destruir las fuerzas estratégicas estadounidenses por medio de un ataque por sorpresa y además intentarán interceptar el mayor número posible de las cabezas nucleares que se hubiesen librado del ataque anterior, antes de que alcancen sus objetivos en la Unión Soviética. El segundo programa está encaminado a reducir hasta unos límites mínimos los daños ocasionados por los proyectiles norteamericanos, estableciendo un plan perfecto de evacuaciones urbanas y de dispersión y defensa de las zonas industriales. Debido a que los rusos podrían "denegar" así importantes objetivos urbanos y económicos ante un posible ataque norteamericano, se argumenta que las fuerzas de los Estados Unidos podrían estar expuestas al fracaso en el logro de los niveles calculados de destrucción asegurada, y que de esta forma Rusia sería capaz de mantener la base industrial, personal y administrativa necesaria para conseguir una rápida recuperación después de la guerra.

A pesar de este supuesto éxito ruso en el empeño, ningún tipo de estrategia podría reducir significativamente los niveles de destrucción de una guerra nuclear generalizada. Más aún, ningún esfuerzo soviético podría salvaguardar y proteger su potencial de recuperación. Por el contrario, tales programas sólo podrían conseguir limitar la capacidad de represalia norteamericana en la perspectiva y dentro de los límites –estrechos y arbitrarios– del concepto de destrucción asegurada tratada más arriba.

Estas elucubraciones parecen ignorar el hecho de que, incluso en las circunstancias más desfavorables, los Estados Unidos estarían capacitados para lanzar un ataque de represalia más que suficiente. Cualquier plan ruso dirigido a disminuir la capacidad de destrucción asegurada norteamericana tendrá que enfrentarse con la formidable tarea de reducir las fuerzas estadounidenses hasta unos niveles muy por debajo de los 400 megatones equivalentes dispuestos para ser lanzados. (En realidad esta tarea podría resultar aún más gigantesca, ya que un ataque norteamericano bien planeado, de alrededor de los 200 megatones equivalentes, podría producir de forma inmediata la muerte de una quinta parte de la población rusa y destruir más de los dos tercios de su industria, satisfaciendo así los requerimientos previstos de destrucción asegurada.)

Resulta del todo improbable que un anticipado ataque inicial ruso pudiera conseguir los efectos deseados. Por una razón fundamental: 400 megatones equivalentes es sólo una fracción del actual arsenal norteamericano. Más de la mitad del actual arsenal de los Estados Unidos –que supera los 6000 megatones equivalentes–, se encuentra repartido entre submarinos lanzadores de misiles estacionados en sus bases, bombarderos en situación de alerta en las bases del Mando Aéreo Estratégico y misiles instalados en silos, todos los cuales pueden entrar en acción a los pocos minutos de la orden del Presidente. El resto de las fuerzas estratégicas norteamericanas consiste fundamentalmente en bombarderos en situación de no alerta y submarinos en sus bases, sometidos a revisiones de mantenimiento. Aun en el caso de que un ataque ruso por sorpresa consiguiese destruir un buen número de misiles norteamericanos con base terrestre en sus silos y todos los bombarderos y submarinos en situación de no alerta, todavía se dispondría de más de 2000 megatones equivalentes para llevar a cabo una acción de represalia. Incluso si por razones desconocidas un número imprevisto de armas norteamericanas no llegasen a funcionar o fuesen destruidas en vuelo, más de 1500 megatones equivalentes podrían aún ser lanzados con un alto grado de fiabilidad. Estas cifras representan las condiciones más desfavorables desde un punto de vista norteamericano; si con anterioridad a un ataque ruso apareciese alguna señal o aviso de alerta, podría ponerse en situación de empleo inmediato un número adicional de bombarderos y submarinos; de esta forma el número de megatones equivalentes dispuestos para ser utilizados sería más del doble.

A causa de la disponibilidad de lo que



CIUDADES (POBLACION)

- DESDE 100.000 A 250.000
- DESDE 250.000 A 1.000.000
- MAS DE UN MILLON

BASES DE MISILES SUBMARINOS

PRINCIPALES AERODROMOS

BASE DE LANZAMIENTO DE MISILES

CENTROS DE PRUEBA DE MISILES

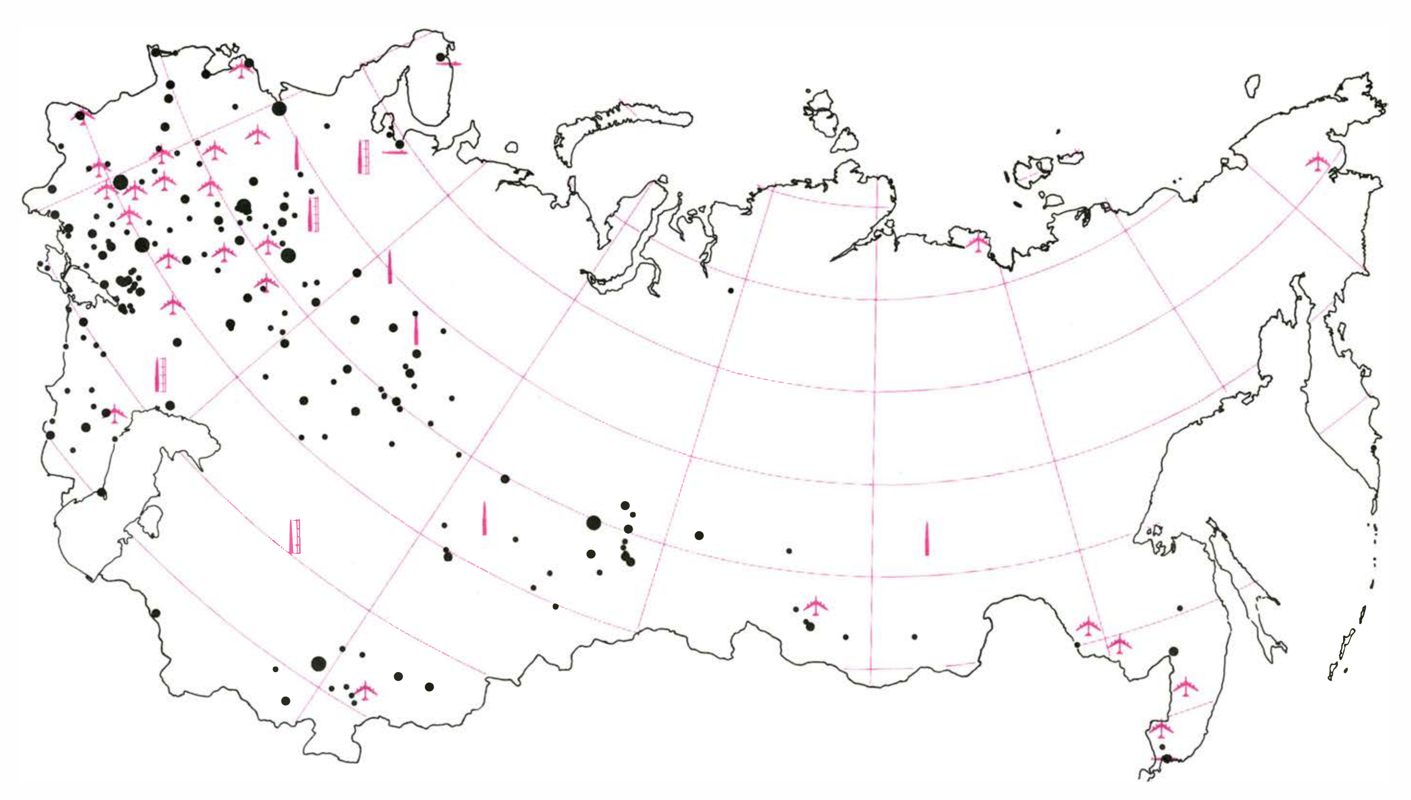
POSIBLES OBJETIVOS DEMOGRAFICOS Y MILITARES en los Estados Unidos y Rusia y su concentración comparada. Los puntos negros indican la localización de los núcleos urbanos más importantes de cada país. Los símbolos coloreados representan las instalaciones de armas estratégicas:

a veces se ha dado en llamar fuerzas de "superdestrucción", cualquier esfuerzo encaminado a reducir el número o la efectividad de las armas nucleares norteamericanas está condenado al fracaso. Por ejemplo, Moscú está bajo la protección de un sistema de misiles antibalísticos (ABM), limitados por tratado a 100 lanzadores de misiles (en la actualidad este sistema cuenta con 64 misiles desplegados). En el caso de que se produjese un ataque de misiles, estas armas rusas podrían destruir cierta cantidad de misiles atacantes. Los planificadores norteamericanos podrían compensar con facilidad esas posibles pérdidas por varios procedimientos, uno de los cuales sería añadir un número adicional de cabezas nucleares al "lote de Moscú", basando dicho incremento en unas previsiones generosas acerca de la efectividad del sistema ABM desplegado en la capital soviética.

El criterio de destrucción asegurada asume de forma implícita la menos probable de las hipótesis sobre una guerra nuclear generalizada. Resulta del todo improbable que una guerra nuclear de estas características pueda dar comienzo por medio de un ataque por sorpresa y masivo de uno de ambos bandos contra las ciudades de su oponente. Una guerra de este tipo, por el contrario, aparecería muy posiblemente como consecuencia de una escalada gradual en la crisis, que podría iniciarse con el lanzamiento de proyectiles nucleares de limitada potencia contra objetivos militares. Estas hipótesis alternativas traen como consecuencia el que las grandes urbes dispondrían de un conjunto muy amplio de medidas de alarma, capaces de detectar la posibilidad o probabilidad de un ataque nuclear, junto con previsiones para la evacuación de la población y la adopción de otras medidas encaminadas a reducir en lo posible toda clase de consecuencias catastróficas. Si los acontecimientos se desarrollasen de este modo, las características normales y habituales de la población afectada podrían haber sufrido un cambio sustancial llegado el momento del ataque. Como la destrucción asegurada se calcula sobre la base de las densidades de población en tiempo de paz y los cálculos se apoyan en determinadas premisas acerca de la disposición de los habitantes y de los trabajadores de un día para otro, los niveles programados de bajas mortales podrían no llegar a alcanzarse en circunstancias reales.

El criterio de destrucción asegurada supone también que una guerra nuclear generalizada pudiera consistir en un único pero brutal "intercambio de mazazos" dirigidos contra objetivos tanto económicos como militares. Estos "golpetazos", sin embargo, podrían muy bien mantenerse controlados durante algún tiempo a un nivel relativamente bajo de intensidad. Un intercambio de golpes limitados podría consistir en ataques contra complejos industriales alejados de los grandes núcleos urbanos, con objeto de desanimar un posible intento por parte del oponente de desarrollar una escalada en sus ataques, ante el riesgo de un elevadísimo número de bajas mortales (los rusos en particular han construido algunas instalaciones clave en zonas remotas, donde podrían ser atacadas con un número de bajas previsibles relativamente escaso). Los ataques contra las ciudades podrían ir precedidos de una advertencia o de un ultimátum que producirían, como es natural, una intensificación en el proceso de evacuación. Incluso una guerra generalizada podría dar comienzo a base de una lenta campaña de "chalaneo" o "intercambio" de ciudades. En otras palabras, cualquier cambio en los supuestos puede alterar de forma sustancial el concepto de destrucción asegurada y hacer surgir dudas muy fundadas sobre la precisión de las previsiones y cálculos de bajas.

En realidad, y con la excepción de algunos casos muy aislados, cualquier país dispondrá siempre de alguna oportunidad de llevar a cabo la evacuación de sus centros urbanos de forma más o menos completa, con independencia de su grado de preparación en una actividad de tal naturaleza (por ejemplo, durante los días 1, 2 y 3 de septiembre de 1939 el Gobierno Británico llevó a cabo



principales bases aéreas, bases de misiles submarinos, bases de lanzamiento de misiles con base terrestre y bases de prueba de misiles (*véase leyenda abajo a la izquierda*). Además de las instalaciones señaladas en estos mapas, los Estados Unidos disponen de una variedad de fuerzas estacionadas por doquier en todo el mundo (en especial en la Isla de Guam y en Alaska). Por lo general, los posibles objetivos están más concentrados en la Unión Soviética.

la evacuación de alrededor de millón y medio de mujeres y niños residentes en las ciudades más importantes de Inglaterra y durante las mismas fechas otros dos millones de personas se trasladaron por su propia iniciativa). La enorme disponibilidad de medios de transporte por automóvil de carácter privado en los Estados Unidos compensará con creces, con toda probabilidad, los planes y el entrenamiento soviético en materia de evacuación.

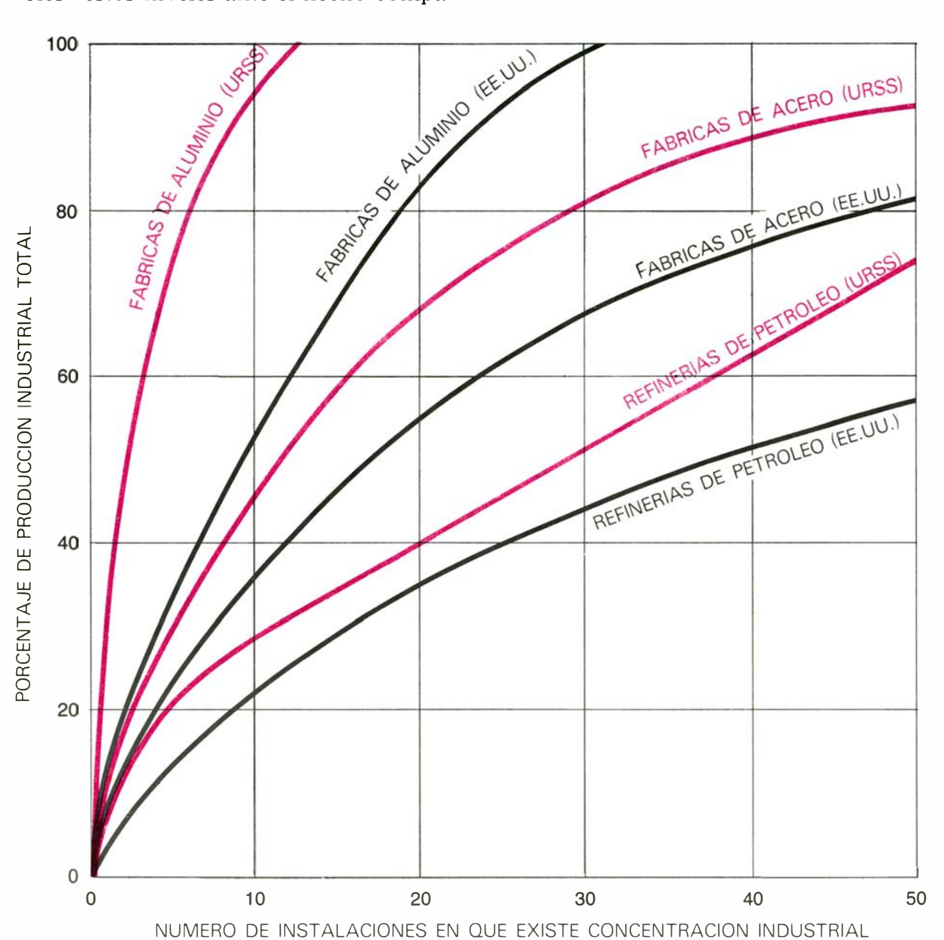
Sin embargo, los comentarios sobre los planes soviéticos de evacuación han despertado muchas preocupaciones en los Estados Unidos. De acuerdo con un reciente informe de la Agencia Central de Inteligencia (CIA), si los rusos llegasen a poder disponer de una semana al menos para llevar a cabo un amplio programa de evacuación de sus ciudades y refugios contra la lluvia radiactiva, las víctimas de la guerra podrían quedar reducidas "a muy pocos centenares de miles, la mitad de las cuales probablemente serían mortales". Algunos analistas han llegado a considerar "aceptables" estos niveles ante el hecho comparativo de que Rusia sufrió la pérdida de 20 millones de muertos durante la Segunda Guerra Mundial. Incluso en la hipótesis de que una evacuación adecuada pudiera reducir el número de bajas inmediatas, la categoría de los daños que Estados unidos podrían ocasionar a las fuentes de la economía rusa, de difícil protección, serían tan grandes que la Unión Soviética quedaría eliminada como futura potencia industrial de primer orden.

Uno de los propósitos de cualquier campaña de bombardeo estratégico es el de reducir las posibilidades enemigas de apoyo a sus fuerzas armadas en sus zonas de operaciones. Durante la Segunda Guerra Mundial, se atacaron con intensidad las fábricas, los sistemas de transporte y las plantas industriales. Entre los objetivos de tales bombardeos estaba la destrucción de las industrias que servían de base a otros sectores de la economía, privándoles de los materiales esenciales para desarrollar su actividad y ocasionando con ello una incapacidad industrial creciente. El bombardeo de Alemania fracasó en la consecución de estos resultados debido, en parte, a las limitaciones impuestas en el tamaño de la carga útil de explosivo químico. Los ataques contra un sistema de objetivos dado tuvo que ser repartido gradualmente a lo largo de varias series de incursiones. De esta forma, las instalaciones supervivientes a las primeras de ellas pudieron ser "reajustadas" para compensar el daño sufrido en ciertos sectores de la red industrial. El personal civil que había quedado sin hogar a causa de los ataques pudo ser cobijado en ciudades cercanas que no habían resultado dañadas. Incluso después del bombardeo atómico de Hiroshima y Nagasaki pudo prestarse una importante ayuda a los supervivientes en los núcleos urbanos próximos a estas dos ciudades.

El despliegue de un gran número de armas nucleares ha modificado de forma radical el concepto de bombardeo estratégico. Las fuerzas habitualmente desplegadas por los Estados Unidos y la Unión Soviética están igualmente capacitadas para lograr la destrucción de la totalidad de las estructuras industriales de cualquier nación. Es más, estos resultados pueden conseguirse de una sola vez, de forma que pueda impedirse la prestación de cualquier auxilio, por pequeño que sea, para dichos objetivos.

Tanto en los Estados Unidos como en la Unión Soviética, el núcleo de la capacidad productiva de muchas industrias clave está concentrado en un limitado número de instalaciones. La economía rusa, de planificación centralizada, se apoya en muchos y vulnerables "cuellos de botella" y puntos de "estrangulamiento". De este modo, la destrucción de un simple objetivo o de muy pocos objetivos puede producir un colapso en muchas otras industrias. A causa de esta concentración, 100 megatones equivalentes, es decir, la carga útil de los misiles transportados por cinco o seis submarinos Poseidón sería suficiente para destruir unas industrias fundamentales, sin las cuales la economía rusa no podrá subsistir por sí misma.

Por ejemplo, un trabajo realizado recientemente bajo la dirección de la Oficina de Asesoramiento Tecnológico del Congreso mostraba que un ataque norteamericano contra las refinerías de petróleo rusas podría conseguir, con sólo 40 cabezas nucleares de escasa potencia, la destrucción de alrededor de las tres cuartas partes del total de la capacidad de refinado rusa. De forma comparativa, con un número pequeño de cabezas nucleares se podría lograr también la destrucción de las fuentes y recursos de transporte, energía, mantenimiento y



OBJETIVOS INDUSTRIALES CLAVE, más concentrados en Rusia que en Norteamérica. En consecuencia, se necesitarían menos cabezas nucleares para dañar la producción soviética de materias de la vital importancia del acero, petróleo y metales no férricos. Además, las economías de ambos países están caracterizadas por peligrosos "cuellos de botella". Por ejemplo, una única planta industrial en Pavlodar, en la Unión Soviética, realiza un trabajo esencial para el 65 por ciento de la industria del aluminio de esta nación. De la misma forma, cerca del 80 por ciento del mineral de hierro embarcado en los Estados Unidos es transportado a través de un conjunto de esclusas sito en Sault Sainte Marie.

administración necesarios para intentar una recuperación económica posterior a una guerra. El sistema energético de la unión Soviética resulta particularmente vulnerable a un ataque y es vital para su posible recuperación. Por ejemplo, la casi totalidad de los transportes de carga entre poblaciones soviéticas se realiza a través de redes ferroviarias electrificadas, mientras que en los Estados Unidos se hace a base de camiones pesados.

Las consideraciones anteriores no pueden llevarnos a la conclusión de que los recursos económicos norteamericanos sean en cierto modo menos vulnerables que los rusos. La Unión Soviética dispone de un número más que suficiente de cabezas nucleares para lograr unos resultados similares sobre los Estados Unidos. Pero resulta instructivo recordar, como hace notar el informe de la CIA, que "la coordinación de las necesidades y las disponibilidades en materia de abastecimiento y transportes constituye un arduo problema para los planificadores soviéticos, incluso en tiempo de paz, por no hablar de los momentos siguientes a un ataque nuclear de alto nivel sobre su territorio".

Incluso en el caso de que los planes de evacuación soviéticos se desarrollasen con éxito, sólo podrían conseguir retrasar –pero de ningún modo impedir– las consecuencias de la guerra sobre su población civil. La base sanitaria, tecnológica y educacional de una nación resultaría igualmente destruida en una guerra nuclear. Tanto las reservas acumuladas, como las instalaciones, podrían ser tomadas también como objetivos. Si lograsen quedar a salvo algunos recursos alimenticios, farmaceúticos, vestidos, equipos o piezas de repuesto, no podría disponerse ni de la estructura administrativa necesaria para la distribución de estos recursos, ni tampoco de los medios de transporte necesarios para hacer llegar estos recursos allí donde fueran precisos. La destrucción de refinerías y centrales eléctricas podría impedir los abastecimientos, y la escasez y falta de recursos aparecería rápidamente por doquier. Los recursos perecederos, incluidos muchos alimentos y medicinas, se desperdiciarían si la energía eléctrica quedase cortada. La destrucción de las viviendas haría difícil la vida en verano e intolerable en invierno. Esta afirmación resultaría particularmente cierta en Rusia donde, aparte de las ciudades, existen muy pocas posibilidades de cobijo como puedan ser, por ejemplo, los hoteles. En pocas palabras, la defensa civil puede proporcionar protección a unos pocos, pero de ninguna forma podrá impedir la destrucción masiva de los recursos esenciales en que se apoya la vida humana. La interdependencia económica dentro de una nación industrializada constituye una vulnerabilidad que no puede ser objeto de una defensa adecuada ni suficiente.

La estrucutura administrativa y social de una nación quedaría también destrozada por un ataque nuclear, hasta el punto de que su sistema político saltaría hecho pedazos, sin posibilidades de reconstrucción. Aunque se están construyendo refugios especiales para proteger el aparato burocrático y la seguridad interna del gobierno soviético ante la eventualidad de una guerra, a los Estados Unidos no les faltarían procedimientos para atacarlos con eficacia.

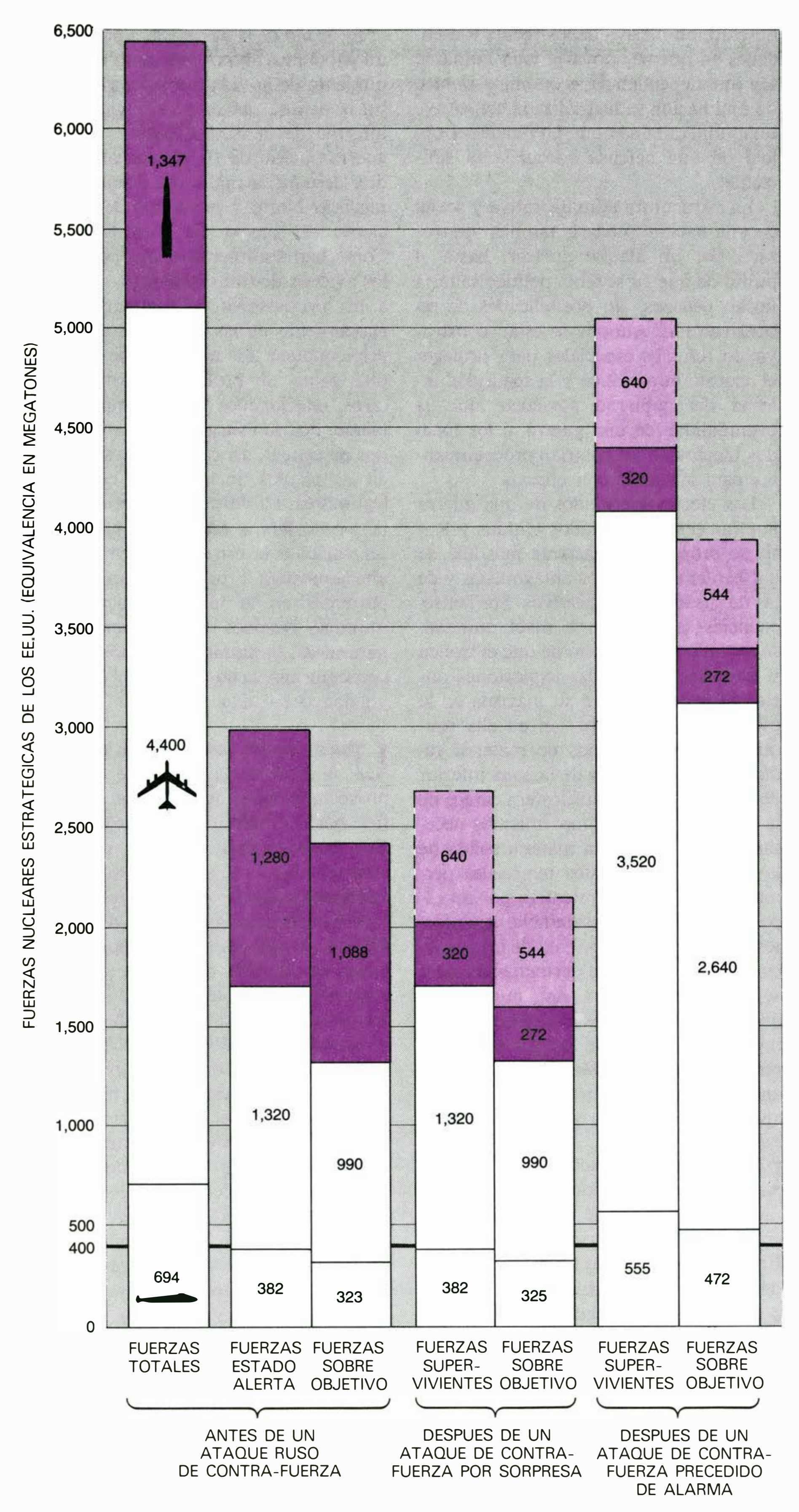
Los efectos retardados de una guerra nuclear entre los Estados Unidos y Rusia se propagarían bastante más allá de las fronteras de ambos antagonistas y de las de sus aliados respectivos. Sus consecuencias se sufrirían a nivel mundial, sobre todo por el hecho de que el tronco y la nube típico de las explosiones nucleares penetrarían en su mayoría en la estratosfera y depositarían en ella residuos de diversos tipos de material radiactivo. A diferencia de la zona inferior de la atmósfera, la estratosfera carece de la humedad y de los movimientos necesarios para separar la materia sólida de la gaseosa. Como estos materiales permanecerían en la estratosfera por un espacio de tiempo considerable, sus efectos se irían diluyendo. Una de las consecuencias de esta larga permanencia sería su amplia dispersión. Así, aunque los efectos en la estratosfera tendrían una menor intensidad que los efectos similares en la baja atmósfera, mostrarían sin embargo una mayor permanencia y abarcarían un mayor radio de acción.

Un informe realizado, en 1975, por la Academia Nacional de Ciencias señalaba tres efectos de la guerra nuclear con un impacto adverso a nivel mundial. En primer lugar, podría agotarse el ozono de la estratosfera, debido a que los óxidos nitrogenados creados por el calor de las explosiones nucleares sobre el oxígeno y nitrógeno de la atmósfera serían inyectados en la estratosfera, donde influirían en la transformación del ozono en oxígeno molecular. En segundo lugar, el depósito de grandes cantidades de polvo y cenizas en la parte alta de la atmósfera podría alterar la cantidad de radiación solar con destino a la superficie terrestre. Por último, peligrosos isótopos radiactivos podrían desparramarse a través de la estratosfera, cayendo lentamente sobre la tierra a escala mundial.

El ozono de la estratosfera desempeña un papel muy importante para el mantenimiento de la vida humana, ya que filtra la dañina radiación ultravioleta. Un informe de la NASA calculó que una guerra nuclear de 10.000 megatones podría destruir la mitad del ozono del hemisferio Norte y alrededor del 30 por ciento del ozono del hemisferio Sur. Como han argumentado los oponentes a los aviones de transporte supersónicos y a los propulsores de fluorcarbono, el agotamiento de las capas de ozono podría conducir a la aparición de una amplia gama de problemas sanitarios y otros relacionados con el medio ambiente. Podrían surgir niveles más elevados de cáncer, así como efectos dañinos en las plantas, incluidos los cereales y legumbres. La destrucción del ozono de la estratosfera a esta escala podría dar un vuelco a la estructura térmica de la alta atmósfera y provocar cambios importantes en la temperatura a escala mundial. Después de una guerra de tal naturaleza, se tardarían muchos años en conseguir que la concentración de ozono volviera a su nivel normal.

Una explosión aislada de un megatón, a nivel de la superficie terrestre, proyectaría miles de toneladas de polvo fino hasta el interior de la estratosfera. Este polvo podría absorber, reflejar y esparcir las radiaciones solares, o reflejadas desde la corteza terrestre; y no han faltado opiniones acerca de la posibilidad de que este hecho pudiera ser causa de cambios profundos en las características atmosféricas de la superficie terrestre. Sin embargo, de acuerdo con el estudio realizado por la NASA, una guerra de 10.000 megatones no llegaría a lanzar sobre la estratosfera más cantidad de polvo que el que fue arrojado por la explosión del volcán Krakatoa el año 1883. Hallando un término medio entre tales tipos de acontecimientos, el informe de la NASA concluye que sólo podrían producirse cambios muy ligeros en las condiciones atmosféricas.

Los isótopos radiactivos serían desparramados a nivel mundial por medio de los procesos de transporte provocados en la estratosfera. Como estos isótopos tendrían un período relativamente largo de permanencia allí, muchos de los de tipo maligno, pero de corta vida, degenerarían antes de que pudieran alcanzar la corteza terrestre. Sin embargo, algunos isótopos de naturaleza lesiva, tales como el estroncio 90, cesio 137, yodo 131 y carbono 14, mantendrían su actividad y podrían introducirse en las cadenas tróficas de la biosfera. El informe de la NASA no hacia mención de la posibili-



ARSENAL ESTRATEGICO DE LOS ESTADOS UNIDOS. Posee unas cantidades de armas nucleares utilizables muy superiores a las necesarias para cumplir la misión de destrucción asegurada, incluso con posterioridad a un eventual ataque masivo ruso por sorpresa sobre sus objetivos militares estratégicos. Si funcionase algún sistema de alarma o de aviso con anterioridad a tal ataque, el número de armas nucleares norteamericanas que podrían ser utilizadas contra objetivos en la Unión Soviética aumentaría considerablemente. La línea negra de trazo grueso que atraviesa el mapa indica la equivalencia de 400 megatones que se considera suficiente para producir la muerte del 35 por ciento.

dad de que esta lluvia radiactiva produjera los efectos letales para la vida humana, a nivel mundial, descritos en novelas tales como *Sobre la playa* de Nivel Shute. Las concentraciones de lluvia radiactiva sobre las naciones combatientes (y sobre las naciones vecinas) podrían representar, sin duda, un riesgo de radiación muy acusado para muchos de los evacuados y de los residentes en zonas rurales que pudieran haber salido bien librados en el ataque a las ciudades. Podrían hacer su aparición "puntos calientes" de menor intensidad a mayores distancias, con consecuencias biológicas adversas. Muy escasas zonas del país atacado se verían libres del peligro representado por la lluvia radiactiva, ya que un ataque de envergadura podría cubrir objetivos económicos y militares a nivel nacional, dejando contaminadas la mayor parte de las zonas.

Los fenómenos atmosféricos son muy complejos y no resulta claro pronosticar cuál sería la influencia de una guerra nuclear de 10.000 megatones en relación con el clima. Aunque del estudio de la NASA se deduce que las consecuencias del agotamiento del ozono y de la carga de polvo no producirían probablemente un impacto irreversible sobre las características climáticas globales, el informe sí indicaba que no podían descartarse cambios de naturaleza mucho más grave. No pueden ocultarse las posibilidades de acciones sinergésicas entre estos posibles efectos. Por ejemplo, se ha hecho notar que un enfriamiento global de un solo grado centígrado podría dar al traste con toda la cosecha de trigo en Canadá. Otras temibles posibilidades incluyen la expansión o la fusión de los hielos polares.

El informe de la NASA no contempla cambios posibles en el clima continental a consecuencia de los efectos ocasionados por causas tales como los fuegos. Una simple explosión en el aire de 10 megatones puede iniciar fuegos en un bosque que acaben cubriendo de ruinas miles de kilómetros cuadrados. La combustión de las inmensas extensiones de prados y bosques de los Estados Unidos y la Unión Soviética podría producir la defoliación de la cubierta natural del terreno, modificando así la superfice de la tierra y dando lugar a cambios en la climatología. Los productos de la combustión de partículas arrojadas al interior de la atmósfera por los incendios producidos en las zonas boscosas absorberán y reflejarán la radiación solar y actuarán a su vez como centros nucleares para la formación de gotitas de agua y cristales

El virus de Epstein-Barr

El más común, quizá, de los virus que infectan a los seres humanos provoca la mononucleosis infecciosa. Se poseen pruebas que lo relacionan con dos tipos de cáncer

Werner Henle, Gertrude Henle y Evelyne T. Lennette

La capacidad que poseen algunos virus de producir tumores en animales, tanto bajo condiciones naturales como experimentales, quedó puesta de manifiesto hace ya más de 70 años. En 1908 V. Ellerman y O. Bang, de Dinamarca, demostraron que la leucemia de las aves podía transmitirse por un virus, y en 1911 Peyton Rous, del Instituto Rockefeller para la investigación médica, reveló que un sarcoma podía transmitirse de forma similar en los pollos. Desde entonces se han venido encontrando numerosos ejemplos de virus causantes de tumores en ratón, rana, conejo, cobaya, vaca, gato, perro e incluso en primates no humanos. Tales hallazgos dieron lugar a una intensa búsqueda de agentes similares en el hombre, basándose en el razonable criterio de que éste no podía quedar exento del azote que afecta a los animales próximos.

Cronológicamente, el primer candidato de virus implicado en cáncer humano es el virus de Epstein-Barr. La fascinante historia de este virus empieza alrededor de 1950, cuando Denis Burkitt, cirujano británico a la sazón en la Universidad de Makerere en Uganda, observó su primer caso de un tumor inusual que más tarde llevaría su nombre. El enfermo era un muchacho joven; presentaba un tumor masivo en la mandíbula que demostró tratarse de un linfoma, esto es, un crecimiento maligno de los tejidos linfoides. El tumor progresaba con notable celeridad, doblando su tamaño de cada 24 a 48 horas. Provocaba una desfiguración grotesca y acababa por causar la muerte al metastizar en otras partes del organismo.

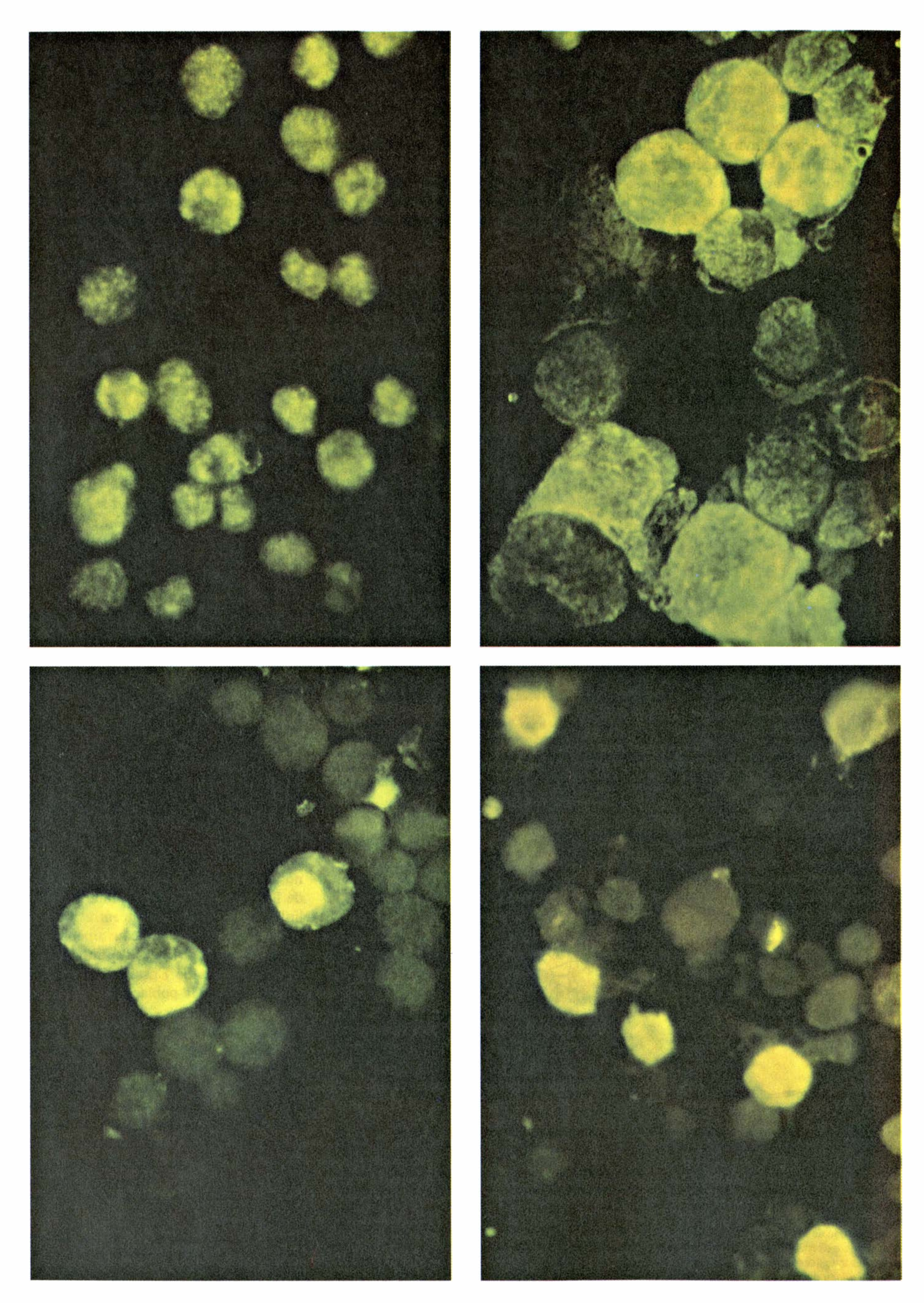
Burkitt se encontró pronto con una excelente oportunidad de observar la enfermedad con mayor extensión. A él llegaron un número sorprendentemente grande de enfermos jóvenes. El linfoma de la mandíbula se presentaba con alta frecuencia en niños de seis a ocho años; era raro en los mayores de 14 y desconocido en los menores de un año. No era infrecuente encontrar grupos de enfermos de ambos sexos en familias vecinas y en un intervalo corto de tiempo. Esta distribución del tumor recordaba las infecciones víricas de la infancia, tales como la poliomielitis antes de la introducción de la vacuna. Burkitt procedió a reunir pruebas epidemiológicas de primera mano acerca del linfoma, y no sólo en Uganda sino por todo Africa. En sus ahora históricos safaris al volante de un land-rover, recorriendo miles de kilómetros, recogió datos sobre los enfermos, su residencia, su familia y sus otras enfermedades. A partir de estos datos, logró hacer un cuadro impresionantemente completo de este cáncer no descrito con anterioridad.

El linfoma de Burkitt fue, y continúa siendo, uno de los tipos de cáncer infantil más común de Africa, con una tasa de un caso por 10.000 niños y año. Se observa sólo en zonas cuya altitud es inferior a los 1800 metros, pluviometría anual superior a los 60 centímetros y temperatura media por encima de los 16 grados Celsius. Estas zonas coinciden con aquellas en que las enfermedades transmitidas por mosquito, tales como la malaria o la fiebre amarilla, son endémicas. Burkitt encontró también pruebas de que el linfoma podía propagarse de una zona a las adyacentes y que ello venía condicionado por variaciones estacionales. Tales hallazgos le indujeron a sugerir, en 1958, que la causa del tumor era un agente infeccioso transmitido por un mosquito, o por lo menos que tal agente era un factor desencadenante de la enfermedad.

Gracias a Burkitt y los médicos que trabajaron con él, se pudo tratar, más tarde, el linfoma letal. Desde los primeros ensayos, el tumor respondía espectacularmente a las inyecciones de ciclofosfamida, disipándose a los pocos días de iniciado el tratamiento. A partir de éste se han desarrollado fármacos más eficaces con efectos colaterales febriles. Sin embargo, en muchos casos el tumor puede eventualmente recurrir y hacerse resistente por último a la quimioterapia, provocando la muerte. Dependiendo de la extensión inicial y localización del tumor, entre el 30 y 60 por ciento de los enfermos sobreviven cinco años o más sin síntoma patológico alguno.

El linfoma de Burkitt pasó a ser objeto de investigación en muchos laboratorios. En 1964, M. Anthony Epstein, B. G. Achong e Y. M. Barr, del Middlesex Hospital de Londres, examinaron al microscopio electrónico células cultivadas a partir de muestras procedentes de biopsias de linfomas de Burkitt. Visualizaron en el interior de las células partículas del virus, que luego recibió el nombre de virus de Epstein-Barr. El virus se clasificó, en razón de su estruc-

LOS ANTIGENOS VIRICOS son proteínas extrañas que inducen respuestas inmunes específicas en los animales y en el hombre. En linfocitos humanos (células sanguíneas blancas) infectados pueden detectarse antígenos específicos contra el virus de Epstein-Barr por la técnica de anticuerpos fluorescentes, como se muestra en las micrografías de fluorescencia de la página opuesta. El descubrimiento de proteínas asociadas a virus en dos tumores humanos ha aportado sólidas pruebas circunstanciales para establecer una relación entre el virus de Epstein-Barr y el cáncer humano. La mayoría de los linfocitos infectados por el virus no producen progenies de partículas víricas, sino que expresan un antígeno vírico en el núcleo celular denominado antígeno nuclear de Epstein-Barr (EBNA, del inglés "Epstein-Barr nuclear antigen"). Cuando los anticuerpos fluorescentes se combinan con este antígeno, hacen que el núcleo celular aparezca brillante, mientras que el citoplasma permanece oscuro (*arriba, izquierda*). Antes de que comience el ciclo de replicación vírica en la célula infectada aparecen los antígenos primarios. El primer antígeno, el componente difuso (*D*), se encuentra tanto en el núcleo como en el citoplasma (*arriba, derecha*). El segundo, el componente restringido (*R*), está limitado a una parte del citoplasma (*abajo, izquierda*). Siguiendo la replicación del ADN vírico, aparecen los antígenos tardíos; estos antígenos son proteínas constitutivas del propio virus. Uno de estos antígenos es el antígeno de la cápside vírica (VCA, del inglés "viral-capsid antigen") que está presente tanto en el núcleo como en el citoplasma (*abajo, derecha*). Cada una de las tres enfermedades asociadas al virus de Epstein-Barr da origen a un patrón distinto de anticuerpos específicos. (Las micrografías de fluorescencia fueron obtenidas por Werner Henle.)



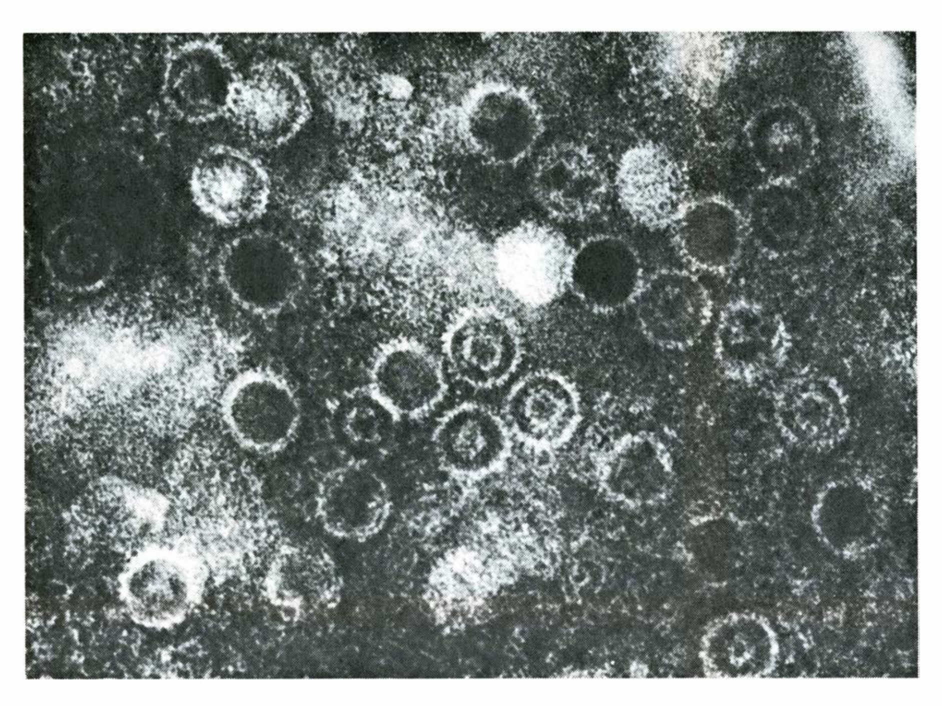
tura, como perteneciente al grupo de los herpes, que contienen su información genética en una doble cadena de ADN. Se conocían en aquel momento tres virus humanos de este tipo: el virus herpes simple, que provoca las "fiebres peribucales"; el citomegalovirus, que induce graves malformaciones congénitas y varios síndromes en enfermos ancianos, y el varicela-zoster, que provoca varicela y afecciones inflamatorias. Pero el descubrimiento de un virus herpes en las células del linfoma de Burkitt no despertó mayor interés, dado que en aquel momento nadie sospechaba que los virus herpes causaran cánceres en animales. Y así, se consideró la presencia de herpes en las células tumorales como un suceso meramente casual.

Mientras que Epstein no lograba recabar ayuda por parte de los virólogos ingleses en la identificación del virus, nosotros probamos suerte con muestras de enfermos de linfoma de Burkitt en Africa para nuestro laboratorio del Hospital Infantil de Philadelphia. Una serie de circunstancias afortunadas determinaron que entráramos en contacto con Epstein; a los pocos días, una colección de sus cultivos celulares cruzaba el Atlántico en dirección a nuestro laboratorio. Poco después de la llegada de los cultivos, pudimos ya demostrar que el virus de Epstein-Barr constituía un nuevo tipo de virus herpes. Llegamos a esta conclusión cuando logramos transmitir el virus a varios cultivos celulares distintos, embriones de pollo y animales de experimentación que eran susceptibles de infección por los tres herpes humanos conocidos.

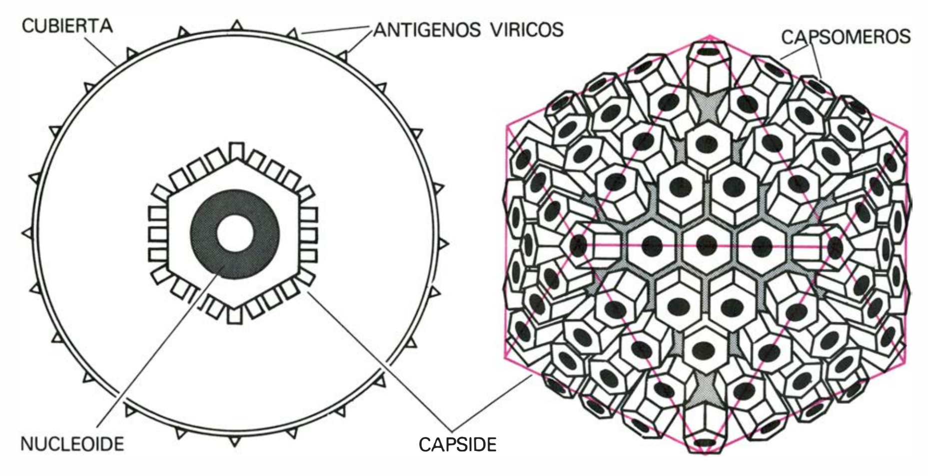
Convencidos de que estábamos ante un virus no conocido con anterioridad, empezamos a investigar la relación entre el virus de Epstein-Barr y el linfoma de Burkitt. Como primer ensayo buscamos anticuerpos específicos del virus en el suero de enfermos. Una vez se ha estado expuesto a un virus, casi invariablemente se fabrican anticuerpos específicos para el mismo, que se desplazan por el torrente circulatorio. La presencia en la circulación de anticuerpos contra un virus determinado en un enfermo es, por consiguiente, un síntoma claro de la exposición anterior al virus.

Para determinar si el virus de Epstein-Barr tenía algo que ver con la génesis del linfoma de Burkitt, nos aprestamos a responder las siguientes preguntas: ¿Presentan todos los enfermos con linfoma de Burkitt anticuerpos contra el virus de Epstein-Barr? ¿Están limitados los anticuerpos contra el virus a los pacientes de esta enfermedad, o pueden presentarse en individuos sanos, o en pacientes con otras enfermedades? Si los anticuerpos no están limitados a los casos de linfoma de Burkitt, ¿presentan los enfermos con linfoma un título (concentración en suero sanguíneo) significativamente elevado, o una variedad de anticuerpos contra el virus de Epstein-Barr) mayor que el resto de la población?

Para detectar los anticuerpos contra virus de Epstein-Barr en enfermos, empleamos una técnica sensible conocida como inmunofluorescencia indirecta. En esta técnica, sobre un portaobjetos de vidrio se fijan células cultivadas que contienen el virus desconocido. Se coloca el suero de un enfermo sospechoso de contener anticuerpos específicos para el virus sobre el portaobjetos; en consecuencia, los anticuerpos del enfermo pueden unirse a los antígenos víricos (proteínas extrañas), ya sea en el interior de la célula o en su superficie. Se



PARTICULAS VIRICAS DE EPSTEIN-BARR aumentadas unas 96.000 veces en esta micrografía electrónica hecha por M. Anthony Epstein. Las cápsides, o proteínas propias, de las partículas víricas se destacan sobre el fondo teñido de oscuro. Algunas contienen en su interior el material de contraste negativo, poniendo de manifiesto si están vacías o poseen el ADN-viral; las partículas víricas no teñidas aparecen blancas. El virus se aisló a partir del linfoma de Burkitt y el carcinoma nasofaríngeo.



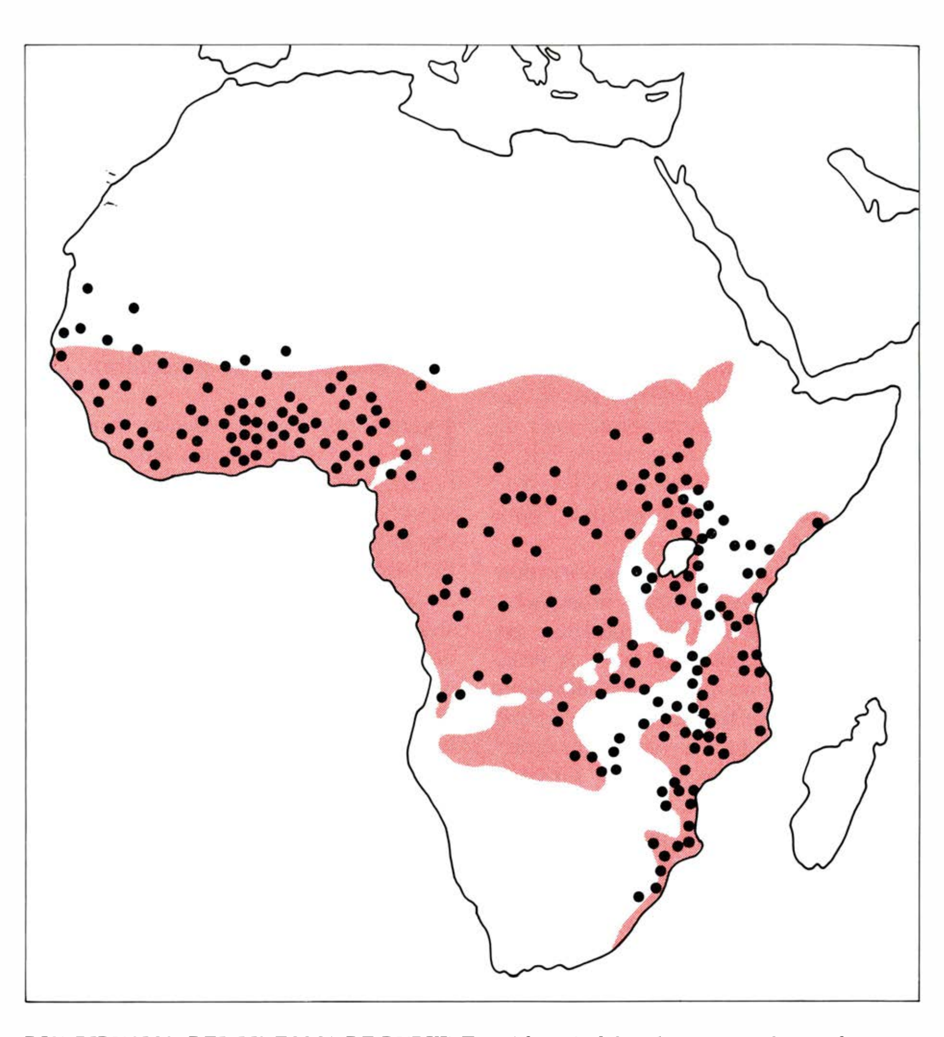
ESTRUCTURA del virus de Epstein-Barr. Tiene la estructura propia del grupo de herpes virus, que poseen su información genética en una doble cadena de ADN. Las partículas víricas infecciosas constan de tres componentes: un núcleo central representado por un círculo, o nucleoide, que contiene el ADN vírico en forma condensada; una cápside que es icosaédrica y está formada por subunidades proteicas tubulares y huecas denominadas capsómeros, y una cubierta de protección que deriva de la membrana nuclear o de la membrana exterior de la célula huésped. La cubierta incorpora varias proteínas víricas (antígenos) que son producidas e insertadas en la membrana de la célula huésped antes de que se inicie el ensamblaje de las partículas víricas completas. Durante la infección, la cubierta vírica se funde con la membrana externa de la célula huésped susceptible, y el nucleocápside (la cápside y el ADN vírico que se encuentra encerrado en su interior) se introducen en la célula. El ADN vírico persiste en estado reprimido o bien desvía el sistema enzimático del huésped hacia la producción de nuevas partículas víricas.

lavan las células; un segundo grupo de anticuerpos que reaccionan específicamente con todos los anticuerpos humanos, y que están unidos a unas moléculas con un marcador fluorescente, se incuban conjuntamente con las células que ya habían incorporado los anticuerpos del enfermo. Cuando se iluminan las células con radiación ultravioleta en un microscopio, los anticuerpos del enfermo se ponen de manifiesto como zonas fluorescentes dentro de las células infectadas con virus. La fluorescencia aparece clara dado que, a cada uno de los anticuerpos del enfermo, pueden unirse varios anticuerpos fluorescentes.

Este sencillo ensayo crítico bastó para confirmar que el virus de Epstein-Barr es un nuevo virus, pues los anticuerpos del resto de los virus herpes humanos conocidos no reaccionan con las células del linfoma de Burkitt. Obtuvimos luego suero de un niño de Nigeria con linfoma de Burkitt que había sido transportado desde Bethesda, para su tratamiento en el Centro Clínico perteneciente a los National Institutes of Health. El test de inmunofluorescencia indirecta en el suero del niño puso de relieve que una baja proporción de células del linfoma de Burkitt brillaban con fluorescencia visible. Cuando las células fluorescentes se examinaron aisladas en el microscopio electrónico, mostraron la presencia de nuevas partículas víricas de Epstein-Barr. Las células no fluorescentes no presentaban partículas víricas.

Para nuestra sorpresa, se encontraron anticuerpos contra el virus de Epstein-Barr no sólo en niños afectados por linfoma de Burkitt, sino también en el suero de casi todos los niños africanos sometidos a ensayo. Más inesperado aún fue el descubrimiento de que niños de otras partes del mundo presentaban anticuerpos; ello implicaba que el virus de Epstein-Barr tenía una distribución universal y que casi nadie escapaba a la infección. Pronto se hizo evidente que la edad a la que se adquieren los anticuerpos contra el virus depende de las condiciones de vida en general. Bajo condiciones de amontonamiento de la población y deficiente higiene, propias de países no desarrollados, todos los niños prácticamente fabrican anticuerpos contra el virus a la edad de cuatro años. En los países desarrollados, la adquisición de anticuerpos contra virus se retrasa con frecuencia hasta la adolescencia o la madurez.

La medición del título de anticuerpos del virus de Epstein-Barr en niños con linfoma de Burkitt reveló que los enfermos presentan un título promedio del orden de ocho a diez veces mayor que los niños sanos. Los enfermos con linfoma, por consiguiente, reaccionaban con el virus de forma anormal. Sin embargo, encontramos que enfermos con algunas alteraciones distintas mostraban también títulos elevados de anticuerpos específicos contra el virus de Epstein-Barr. Principalmente entre los enfermos de carcinoma nasofaríngeo, con un título de anticuerpos diez veces superior al de individuos sanos o enfermos de otros tumores de cabeza y garganta. La conexión entre este carcinoma y el virus de Epstein-Barr fue propuesta inicialmente por Lloyd F. Old y sus colegas, del Memorial Sloan-Kettering Cancer Center; éstos observaron que el suero procedente de enfermos que tenían carcinoma nasofaríngeo formaba precipitados con extractos de células cultivadas del linfoma de Burkitt. Tal hallazgo implicaba que el virus de Epstein-Barr estaba en íntima relación con ambas enfermedades.



DISTRIBUCION DEL LINFOMA DE BURKITT en Africa. Tal distribución pone de manifiesto que la mayoría de casos (*negro*) se encuentran en áreas donde la malaria es endémica, como consecuencia de factores climáticos (*color*). El linfoma de Burkitt es un tumor maxilar y de otros tejidos que se presenta en niños africanos y que está estrechamente relacionado con el virus de Epstein-Barr. Este dato sugiere que la malaria puede ser un factor secundario en la génesis del linfoma. (Datos de Denis Burkitt y otros.)

El carcinoma nasofaríngeo es un tumor maligno que se presenta en el revestimiento epitelial del espacio posterior de la nariz. Afecta principalmente a adultos y se presenta con una frecuencia elevada (10 por 100.000 y año) en China meridional y a unas tasas algo menores en Túnez y Africa oriental. Los esquimales norteamericanos que viven en Alaska padecen también esta enfermedad, que es rara en otras partes del mundo. El tumor suele seguir uno o dos cursos letales: puede desarrollarse a través de los huesos del cráneo para invadir el cerebro, o bien las células cancerosas pueden extenderse a partir del tumor primario a través de los vasos linfáticos, formando metástasis en los nódulos linfáticos del cuello y luego en el hígado, pulmones, otros órganos y huesos.

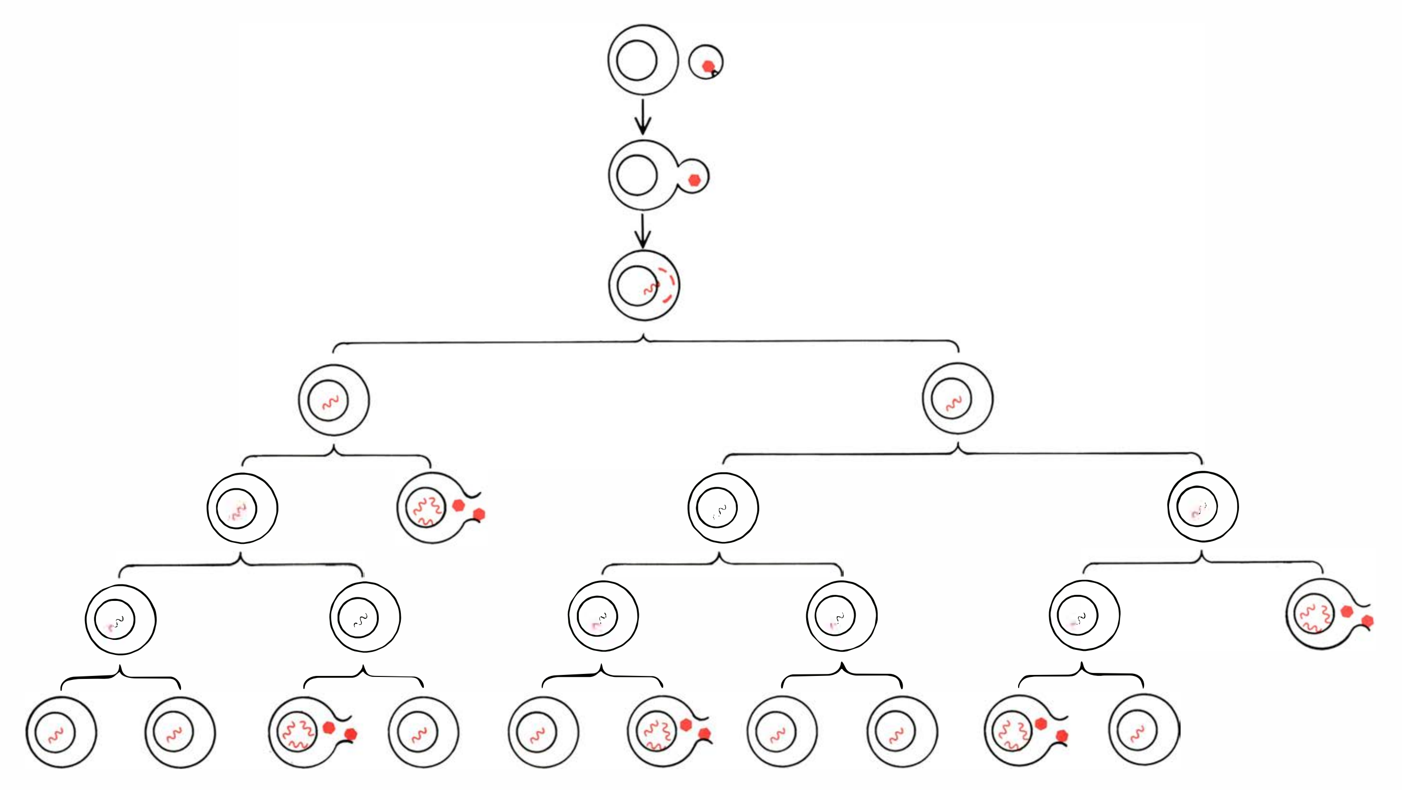
La omnipresencia de anticuerpos contra el virus de Epstein-Barr en poblaciones de individuos sanos nos llevó a sospechar que el virus constituyó, en un comienzo, la causa de una enfermedad común autolimitada, además de estar aparentemente asociado a dos alteraciones malignas. Por una gran coincidencia, la identidad de esta enfermedad se conoció a través de un descubrimiento

accidental en nuestro laboratorio. A finales de 1967, uno de nuestros técnicos jóvenes desarrolló los síntomas clásicos de la mononucleosis infecciosa: dolor de garganta, fiebre y engrosamiento de los nódulos linfáticos del cuello. No había presentado antes del episodio anticuerpos específicos contra el virus de Epstein-Barr, pero con el inicio de la enfermedad aparecieron anticuerpos en su suero. Este acontecimiento fortuito proporcionaba el primer indicio de que el virus de Epstein-Barr podía ser la tan buscada causa de la mononucleosis infecciosa descrita por primera vez por T. P. Sprunt y F. A. Evans, de la Facultad de Medicina de la Universidad Johns Hopkins, en 1920.

En su forma clásica, la mononucleosis infecciosa provoca fiebre, amigdalitis, hinchazón de los nódulos linfáticos y engrosamiento de hígado y bazo. A pesar de su nombre no parece ser altamente infecciosa; no se han descrito epidemias de importancia. La mayoría de los enfermos son jóvenes; en los casos sin complicaciones, sigue un curso de una a cuatro semanas con una recuperación completa. Los enfermos que presentan complicaciones (encefalitis, neumonía y fallo renal) no llegan al uno por ciento y los casos fatales son muy raros. Sin embargo, algunos clínicos han descrito la mononucleosis como una leucemia autolimitada. Los enfermos muestran un considerable incremento del número de glóbulos blancos, que además son anormales desde el punto de vista morfológico. La proliferación de glóbulos blancos se extiende a los linfocitos B (que se originan en la médula ósea y fabrican anticuerpos) y los linfocitos T (que maduran en el timo y participan en las reacciones de inmunidad celular). En enfermos de mononucleosis, la proliferación anormal de linfocitos desciende pronto; en la leucemia, las células continúan multiplicándose sin control.

Para confirmar una relación causal entre el virus de Epstein-Barr y la mononucleosis, necesitábamos muestras de suero procedentes de pacientes con dicha enfermedad. Recordamos que James C. Niederman y Robert W. McCollum, de la Facultad de Medicina de la Universidad de Yale, se habían comprometido desde 1958 en un estudio prospectivo de la mononucleosis entre los estudiantes de la facultad recogiendo suero de cientos de estudiantes antes (a la entrada en la facultad), durante y después de la enfermedad. Sus valiosas colecciones fueron generosamente puestas a nuestra disposición para realizar ensayos. Encontramos que los anticuerpos contra el virus de Epstein-Barr siempre estaban ausentes antes de la enfermedad, pero se presentaban durante y después de la afección. Estos hallazgos nos llevaron a la conclusión inequívoca de que el virus de Epstein-Barr es responsable de al menos una enfermedad: la mononucleosis infecciosa. Estudios epidemiológicos subsiguientes han indicado que la población que ha tenido mononucleosis infecciosa no presenta una elevada probabilidad de contraer posteriormente un cáncer. Este criterio concuerda con el hecho de que virtualmente todo el mundo se infecta con el virus de Epstein-Barr, incluso los que no desarrollan mononucleosis.

Se han hecho grandes progresos en el esclarecimiento de cómo el virus de Epstein-Barr provoca la mononucleosis. Se sabe ahora que la ruta de infección a individuos no inmunes es a través de la boca y el tracto respiratorio, ya sea por el uso de cubiertos, platos o vasos contaminados, contacto oral íntimo (en el lenguaje coloquial se la denomina "enfermedad del beso") o la inhalación de partículas víricas. El primer blanco celular para los virus lo constituyen probablemente las células epiteliales de la boca y la garganta; pero antes o después el virus infecta los linfocitos B, que lo diseminan a través del sistema linfático (por alguna razón, el virus de Epstein-Barr es altamente selectivo para los linfocitos B y no infecta a los T). Los primeros síntomas de la enfermedad suelen desarrollarse en un período de cuatro a siete semanas después de la infección; durante la fase aguda de la enfermedad pueden detectarse partículas víricas en la saliva de casi todos los enfermos.



INDUCCION ESPONTANEA de la replicación del virus de Epstein-Barr. La inducción puede tener lugar en los descendientes de una sola célula procedente de un cultivo "productor" de células de linfoma de Burkitt. Muchas células que contienen el genoma vírico (la dotación completa de los genes víricos) persisten indefinidamente dividiéndose cada 18-24 horas. Pero en más del 10 por ciento de las células, el genoma vírico reprimido se activa y conduce a la elaboración de partículas víricas. La liberación de progenies víricas va acompañada de la muerte y lisis de la célula huésped.

En el curso de su evolución, el virus de Epstein-Barr ha establecido una simbiosis casi perfecta con su huésped humano, asegurando la supervivencia de ambos. Si bien la infección produce una inmunidad permanente, el virus persiste en un estado latente en los linfocitos B del sistema linfático y se activa de forma espontánea en algunas células. Periódicamente son liberados en la saliva progenies de partículas víricas infecciosas, que aseguran la llegada del virus a los individuos no inmunes.

Bajo condiciones primitivas, tales como las prevalecientes en los tiempos prehistóricos y aún hoy en muchas partes del mundo, las infecciones del virus de Epstein-Barr ocurren en la primera edad; entonces no producen ninguna enfermedad significativa debido quizás a la inmadurez de las células blanco o del sistema inmune. En un ambiente privilegiado con mejores condiciones higiénicas y un modo de vida propio del mundo desarrollado, las infecciones primarias por el virus se retrasan hasta el final de la infancia e incluso de la adolescencia, cuando ocurre que es más probable que provoquen los síntomas de la mononucleosis. Por ejemplo, en poblaciones escolares de Estados Unidos, entre el 35 y el 80 por ciento de los estudiantes al incorporarse tienen anticuerpos contra el virus de Epstein-Barr; del orden del 10 al 15 por ciento de los que no poseen anticuerpos acusan la infección cada año. Cerca de los dos tercios de estos individuos infectados tardíamente manifiestan alteraciones significativas. Esto explica por qué la mononucleosis es esencialmente una enfermedad propia de los países económicamente avanzados.

El virus de Epstein-Barr es similar a algunos de los demás herpes en el sentido de que puede permanecer "oculto" en un estado silencioso, o reprimido, en células linfoides humanas en división, sin inducir la fabricación de progenie vírica ni la muerte de la célula huésped. Los linfocitos B humanos cultivados, infectados y con el virus "latente", pueden clasificarse en dos grupos, a saber: líneas celulares productoras y líneas celulares no productoras. Líneas celulares productoras son aquellas en las cuales el genoma vírico (todos los genes víricos) pueden activarse de una forma espontánea en una proporción baja de células (superior al 10 por ciento), dando origen a antígenos víricos y ocasionalmente a partículas víricas infecciosas. Líneas celulares no productoras son aquellas en las que el genoma vírico no se activa de una forma espontánea. E incluso cuando el genoma vírico de una célula no productora se activa por vía experimental, el proceso suele ser abortivo: se fabrican antígenos pero no partículas víricas.

La presencia de genomas víricos en células no productoras quedó demostrada por primera vez en 1970 por Harald zur Hausen y Heinrich Schulte-Holthausen, de la Universidad de Erlangen. Mostraron que las cadenas de ADN vírico, extraídas a partir de partículas víricas infecciosas, se apareaban con cadenas complementarias de ADN extraídas de células no productoras. Observaron que las células con infección latente contenían unas 50 copias completas del genoma vírico.

Alice Adams y Thomas Lindahl, del laboratorio de George Klein del Instituto Karolinska de Estocolmo, descubrieron que sólo parte del ADN vírico podía separarse de la célula huésped por centrifugación. La fracción separable está formada por plásmidos de ADN vírico celular, que son de la misma longitud que las hebras de ADN incluidas en la partícula vírica infecciosa. La fracción no separable está representada por copias de ADN vírico que se han insertado en forma lineal en el ADN de la célula huésped, donde se replican con el resto de los cromosomas del hospedante.

La represión del ADN vírico puede revertir en una célula de una línea productora, conduciendo a la producción activa de partículas víricas y a la muerte celular. Las partículas víricas liberadas al medio de cultivo, a partir de estas células productoras, son de dos tipos: transformantes y líticas. Los virus líticos inducen la fabricación de partículas víricas hijas en las células nuevamente infectadas. Los virus transformantes, que son liberados por la mayoría de las células productoras, provocan una multiplicación indefinida de los linfocitos B cuando se les hace crecer en cultivo de tejidos. (Los linfocitos B no infectados son incapaces de crecer y dividirse en cultivos, y acaban por morir a las dos o tres semanas.) Si se inactiva el virus transformante por calentamiento o tratamiento con anticuerpos, o se elimina del medio de cultivo por paso de éste a través de un filtro extremadamente fino, los linfocitos B no se transforman. Estos experimentos han establecido con nitidez que el agente transformante es el propio virus.

El virus de Epstein-Barr transforma también linfocitos del ser humano vivo. Los linfocitos B obtenidos de la sangre y los nódulos linfáticos de individuos infectados por el virus crecen fácilmente en cultivos celulares, mientras que no lo hacen células similares procedentes de individuos no infectados. Giovanni Roc-



MOLECULA DE ADN que porta la información genética del virus de Epstein-Barr. Se presenta en múltiples copias tanto en linfocitos con infección latente como virulenta. La mayoría de las moléculas de ADN adquieren una forma circular, como aquí se muestra y se replican independientemente de los cromosomas de la célula huésped. El resto de las moléculas de ADN vírico están integradas en el ADN de la célula huésped en forma lineal y se replican con él. Dentro de la partícula vírica, el ADN se encuentra en la llamada forma lineal. La micrografía fue realizada por Gunnar Bjursell, de la Universidad danesa de Aarhus.

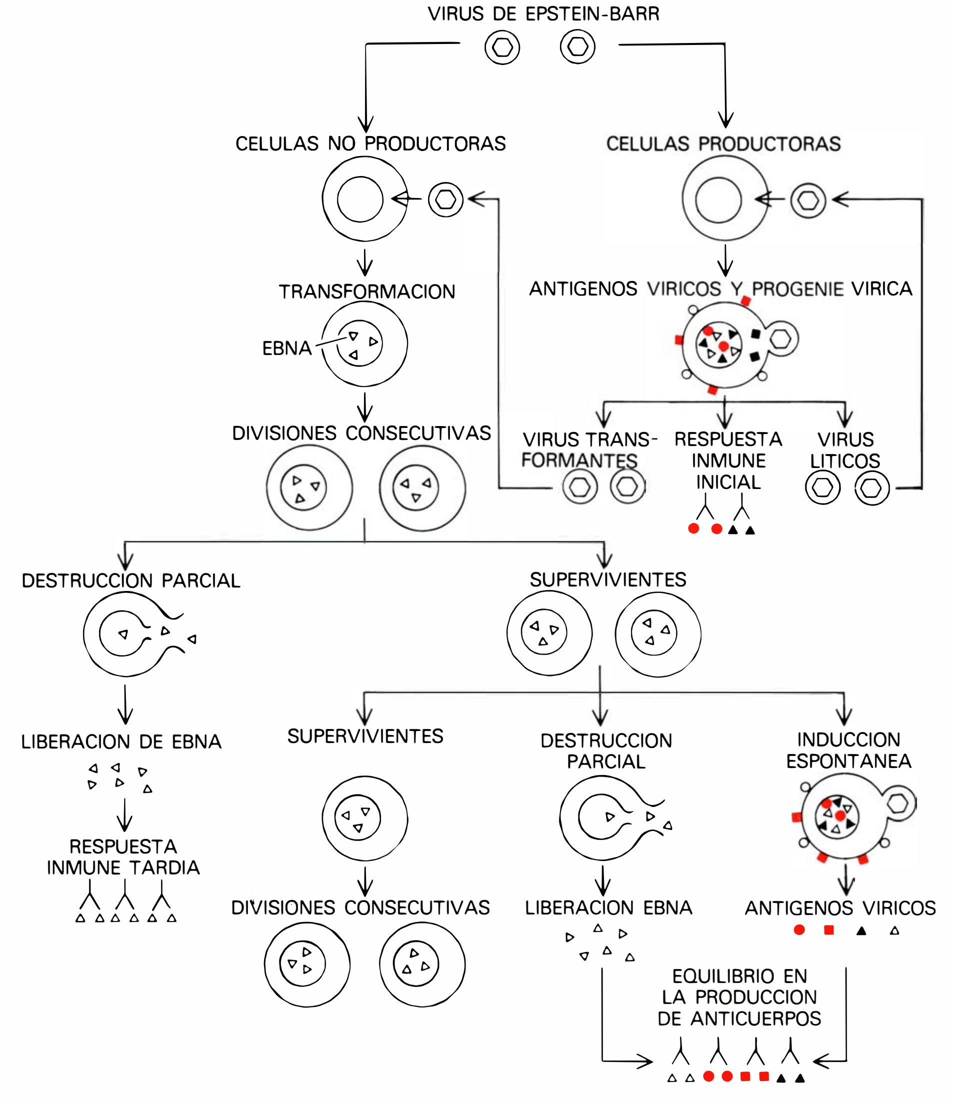
chi y sus colegas, de la Universidad de Roma, han mostrado que, en poblaciones sanas con infección latente del virus de Epstein-Barr, hay un linfocito de cada 10 millones transformado; en enfermos con mononucleosis infecciosa, uno de cada 5000 está transformado. Estos resultados sugieren que el virus de Epstein-Barr es potencialmente inductor de tumores, porque confiere a los linfocitos B la capacidad para un crecimiento y proliferación sin restricciones, propiedad característica de las células cancerosas. La transformación de los linfocitos parece venir inducida por copias de ADN vírico que se han insertado linealmente en el cromosoma del huésped. Elliot Kieff, de la Facultad de Medicina de la Universidad de Chicago, está analizando actualmente el ADN vírico en un esfuerzo por identificar el segmento específico de ADN que codifica para la actividad transformante del virus.

Aun cuando muchos de los ADN víricos de las células no productoras están reprimidos, como señala la ausencia de la replicación vírica, una pequeña fracción del ADN vírico se expresa en forma de proteínas nuevas en el núcleo de la célula y en la membrana celular. Estas proteínas no son estructurales del virus transformante, aunque están codificadas por él. Dado que las proteínas asociadas al virus difieren de las proteínas celulares normales, inducen la fabricación de anticuerpos. Uno de estos antígenos es el antígeno nuclear de Epstein-Barr (EBNA), que se localiza en el núcleo de la célula asociado con la cromatina (material que incorpora el ADN). Descubierto por Beverly Reedman y Klein, quienes mostraron posteriormente que el EBNA se expresa en todas las células que portan el genoma vírico, entre las que se cuentan las células del linfoma de Burkitt y del carcinoma nasofaríngeo. En células no productoras, la presencia de EBNA en el núcleo es el único indicio de que las células han sido transformadas por el virus de Epstein-Barr.

Por manipulación de las condiciones de crecimiento en cultivos celulares de linfocitos con infección latente, o por tratamiento con ciertos fármacos, podemos inducir que células no productoras acometan el ciclo lítico que culmina con la muerte y lisis (disolución) de la célula. El ciclo lítico implica la fabricación de algunos antígenos más y, si avanza lo suficiente, la liberación de la progenie de partículas víricas. Poco después de la inducción del ciclo lítico, y antes de que pueda detectarse cualquier replicación de ADN vírico, aparecen en las células tres antígenos víricos específicos. Estos antígenos se denominaron complejo antigénico primario. El primer antígeno, el componente difuso (*D*), aparece inicialmente en el núcleo celular y se extiende a través del citoplasma. El segundo antígeno, el componente restringido (*R*), se localiza en grandes cantidades en el citoplasma adyacente al núcleo celular. El tercer antígeno, el primero de los de membrana, se inserta en la membrana de la célula huésped. En muchas células el proceso infeccioso se detiene aquí.

Sin embargo, en otras células, el ciclo lítico continúa con la replicación del ADN vírico. A continuación de esta fase aparecen los antígenos víricos tardíos: las proteínas que son componentes estructurales de la cápside o cubierta proteica de la partícula vírica. Los antígenos tardíos incluyen el antígeno de la cápside del virus (VCA) y los antígenos de membrana tardíos, insertos en la membrana celular. Los viriones que forman la descendencia se ensamblan en el núcleo celular; algunos pasan después al citoplasma de la célula y salen extruidos a través de la membrana celular. En el proceso, la cápside vírica se reviste de una cubierta protectora formada a partir de la membrana de la célula huésped. La liberación de partículas víricas hijas –en realidad la síntesis de antígenos– conduce inevitablemente a la muerte de la célula.

La secuencia de sucesos siguientes a la infección con el virus de Epstein-Barr



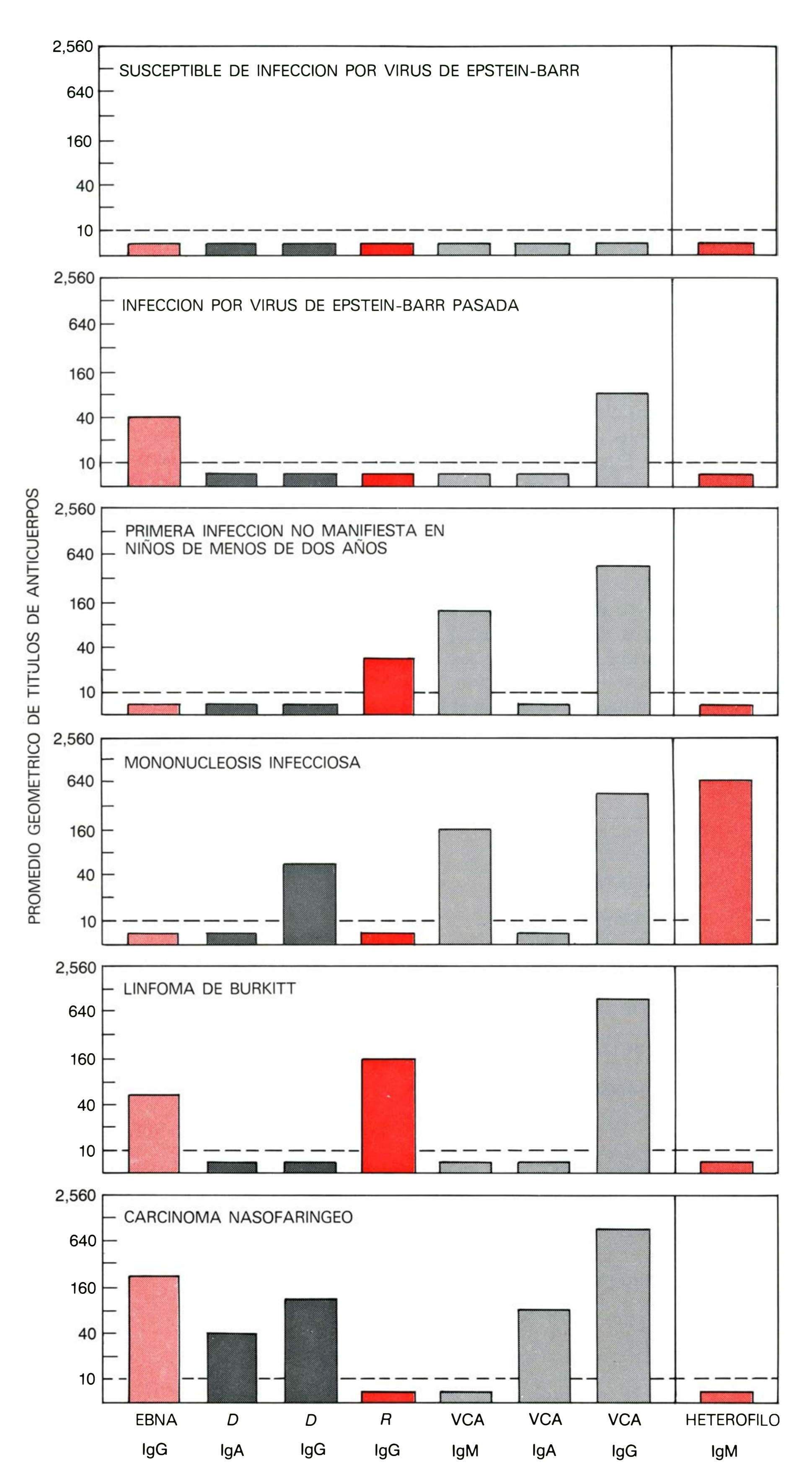
SECUENCIA DE FENOMENOS durante la mononucleosis infecciosa, resumida en este gráfico. Muchos linfocitos B experimentan una infección lítica por el virus de Epstein-Barr, como infección progresiva. Fabrican antígenos víricos y partículas víricas antes de su muerte y lisis inevitable. Los antígenos víricos inducen una respuesta inmunitaria inicial. La progenie vírica está formada en su mayor parte por virus del tipo transformante; los linfocitos B infectados por éstos permanecen vivos y se dividen indefinidamente. Las células transformadas encierran copias de ADN vírico y poseen el antígeno nuclear de Epstein-Barr y un antígeno de la superficie celular que se convierte en objetivo de los linfocitos T activos. La destrucción consiguiente de los linfocitos B transformados por efecto de los linfocitos T provoca algunos de los síntomas clínicos de la mononucleosis; libera también suficiente EBNA para inducir una respuesta inmune tardía a este antígeno. Las células transformadas nunca se eliminan del todo. Algunas de las supervivientes son inducidas a elaborar partículas víricas. Otras son eliminadas de forma continua, liberando antígenos y originando un nivel equilibrado de anticuerpos.

en el hombre es mucho menos clara, porque el período de incubación entre la exposición y la aparición de los síntomas clínicos de la mononucleosis dura de unos 30 a 50 días. Una forma de seguir el curso de la infección es controlar la respuesta inmune del enfermo. En la infección vírica inicial, el primer anticuerpo que aparece tras la infección es la inmunoglobulina M (IgM), que es específica para el VCA. Durante las cuatro o siete semanas del período de incubación, se detiene la producción de IgM y es substituida por la producción de inmunoglobulina G (IgG). Y cuando el enfermo recibe atención médica, los anticuerpos IgM e IgG específicos para VCA han alcanzado, normalmente, sus niveles máximos.

Más del 80 por ciento de los enfermos de mononucleosis desarrollan también anticuerpos IgG transitorios y específicos para el componente *D* del complejo antigénico primario, contemporáneamente al ataque de la enfermedad. Por contra, los anticuerpos para EBNA se desarrollan en una fase adelantada del curso de la misma y persisten toda la vida.

Los enfermos de mononucleosis infecciosa desarrollan asimismo anticuerpos heterófilos, llamados así porque son capaces de dar reacciones cruzadas con antígenos no relacionados procedentes de especies animales distintas. Por ejemplo, los anticuerpos humanos heterófilos aglutinarán glóbulos rojos de cordero o de caballo. Los anticuerpos heterófilos fueron descubiertos en 1932 por John R. Paul y W. W. Bunnell, de la Facultad de Medicina de Yale. Se vio que eran específicos de la mononucleosis. La detección de estos anticuerpos en el suero de enfermos constituye todavía el método diagnóstico elegido para la mononucleosis, en razón de la simplicidad del ensayo. Pero, según parece, el antígeno responsable del anticuerpo heterófilo no está directamente relacionado con el virus de Epstein-Barr, y no todos los enfermos desarrollan estos anticuerpos.

Como hemos indicado, los niños de países del Tercer Mundo se exponen al virus de Epstein-Barr en las primeras etapas de la vida. Estudios recientes en niños de Ghana realizados por nosotros en colaboración con Robert J. Biggar, de los National Institutes of Health, han puesto de manifiesto que los anticuerpos maternos pasan a través de la placenta materna al feto, y protegen al recién nacido de la infección hasta unos seis meses después del nacimiento. En los doce meses siguientes, el 80 por ciento de los niños estudiados desarrollan sus propios anticuerpos contra el virus, sin



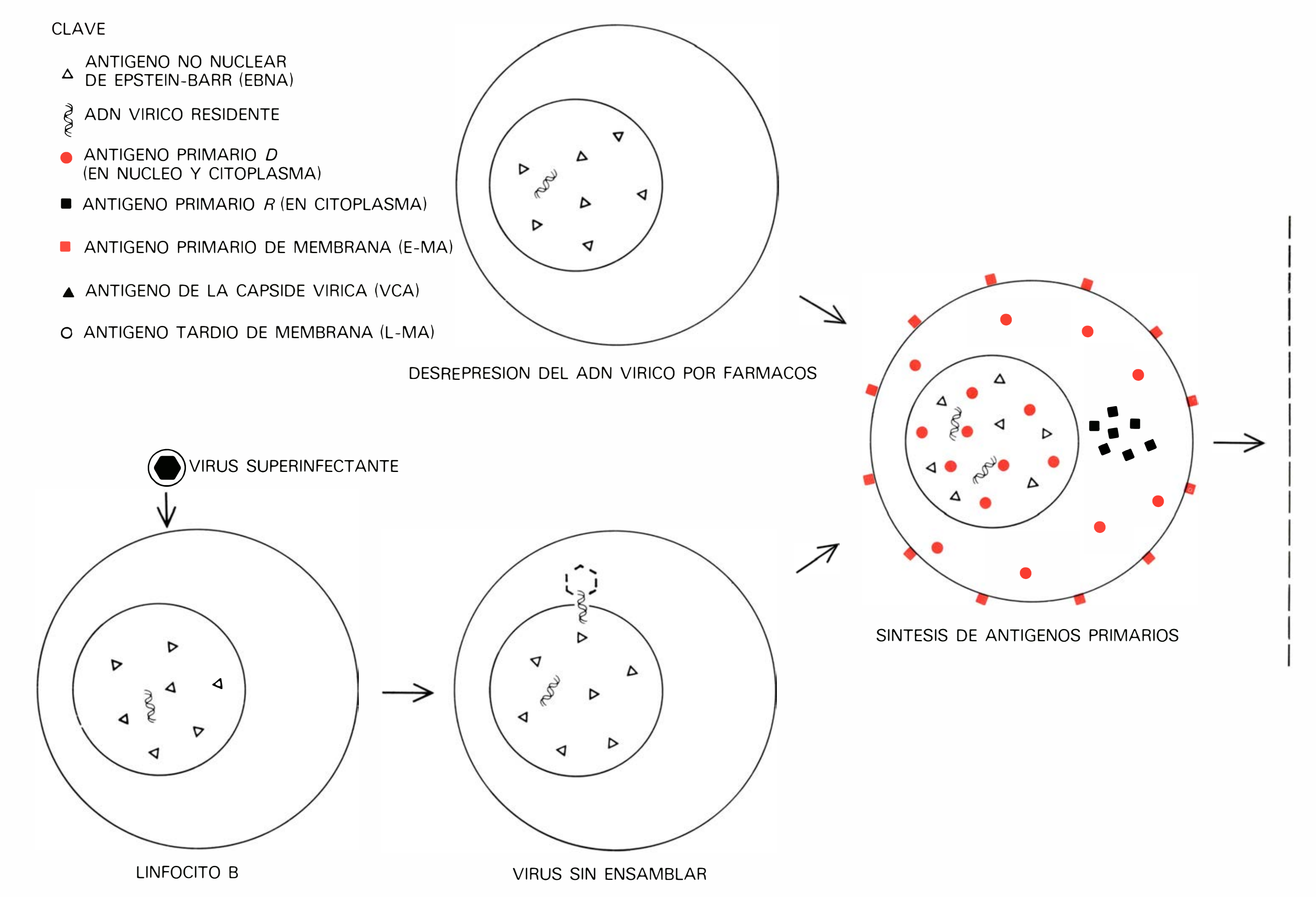
LA RESPUESTA de elaboración de anticuerpos al virus de Epstein-Barr difiere significativamente en las distintas enfermedades asociadas con el virus. Se han identificado los patrones de respuesta característicos para la primera infección "silenciosa" en niños de menos de dos años, la mononucleosis infecciosa en adultos jóvenes, el linfoma de Burkitt y el carcinoma nasofaríngeo. Los enfermos pueden elaborar anticuerpos de tres clases (IgM, o inmunoglobulina M, IgA, o inmunoglobulina A, IgG, o inmunoglobulina G) que actúan sobre antígenos relacionados con el virus: el antígeno nuclear de Epstein-Barr, el antígeno de la cápside vírica y los componentes difuso y limitado (*D* y *R*) del complejo antigénico primario. Los anticuerpos contra el antígeno heterófilo se presentan solamente en enfermos con mononucleosis.

que se manifieste ningún síntoma clínico de enfermedad. La respuesta de los niños al virus de Epstein-Barr es análoga a las respuestas observadas en adultos afectos de mononucleosis, con dos excepciones sorprendentes: los anticuerpos heterófilos están ausentes, o sólo alcanzan niveles muy poco significativos, y rápidamente los anticuerpos IgG contra el complejo antigénico primario se dirigen al componente *R*, no contra el componente *D* (los antígenos específicos para el componente *R* han resultado característicos de enfermos con linfoma de Burkitt). No se ha dilucidado cuál sea el significado de los distintos niveles de anticuerpos en niños y adultos, infectados con el virus de Epstein-Barr.

Mientras las pruebas que relacionan el virus de Epstein-Barr con la mononucleosis no admiten objeción, la implicación del virus en los dos tipos de cáncer está un poco en el aire. Hasta ahora, el argumento en pro de una relación causal entre el virus de Epstein-Barr y el linfoma de Burkitt se apoya en las siguientes pruebas: (1) la detección de ADN vírico o antígenos víricos en el tumor, (2) la transformación por el virus de linfocitos B cultivados, (3) la inducción de linfomas por el virus en primates no humanos, (4) la demostración de títulos elevados de anticuerpos específicos en enfermos de linfoma comparado con los títulos de una población control y (5) la correlación entre los títulos de anticuerpo y el pronóstico del tumor. Estas cinco razones aportan una base bastante bien fundada a la hipótesis de una relación causal entre el virus de Epstein-Barr y el linfoma de Burkitt, aunque siguen sin respuesta cuestiones importantes.

El ADN del virus de Epstein-Barr se ha encontrado en la gran mayoría (98 por ciento) de biopsias de linfomas de Burkitt en Africa, pero en ningún otro tipo de linfoma. Cada célula tumoral contiene múltiples copias del genoma vírico. En las células con infección lítica que incorporan los primeros antígenos, no se detecta ni VCA ni progenie vírica en biopsias recientes de células tumorales, aunque sí aparecen al siguiente o a los dos días del cultivo celular. Es probable que las células con infección lítica aumenten en poca cantidad dentro del tumor, pero sean rápidamente destruidas por las defensas inmunitarias del enfermo tan pronto como aparezca el primer antígeno de membrana sobre la superficie celular. Realmente, ya en 1966 Klein y sus colegas habían detectado un antígeno en la superficie de las células del linfoma de Burkitt, hallando que se encontraba cubierto con anticuerpos propios del enfermo. Por consiguiente, el virus de Epstein-Barr está íntimamente asociado a las células del linfoma de Burkitt.

Sin embargo, cultivos de linfocitos B transformados por el virus de Epstein-Barr difieren en aspectos importantes respecto de cultivos de células del linfoma de Burkitt, que, a su vez, derivan de linfocitos B. Mientras que los cultivos de linfocitos transformados crecen a partir de diferentes células, los cultivos de linfomas de Burkitt son monoclonales; cada tumor crece a partir de una única célula maligna transformada. Además, la gran mayoría de las células del linfoma de Burkitt muestran una anormalidad cromosómica que no se encuentra en los linfocitos transformados ni en los linfocitos de enfermos con mononucleosis. Esta anomalía, observada por vez primera por G. Malonov, Y. Malonova y sus colegas, de la Universi-



CICLO LITICO de replicación del virus de Epstein-Barr. El ciclo puede inducirse en un cultivo de células de un linfoma de Burkitt por superinfección por el virus o por desrepresión del genoma vírico residente por efecto de fármacos inhibidores del metabolismo. Cuando comienza la síntesis del

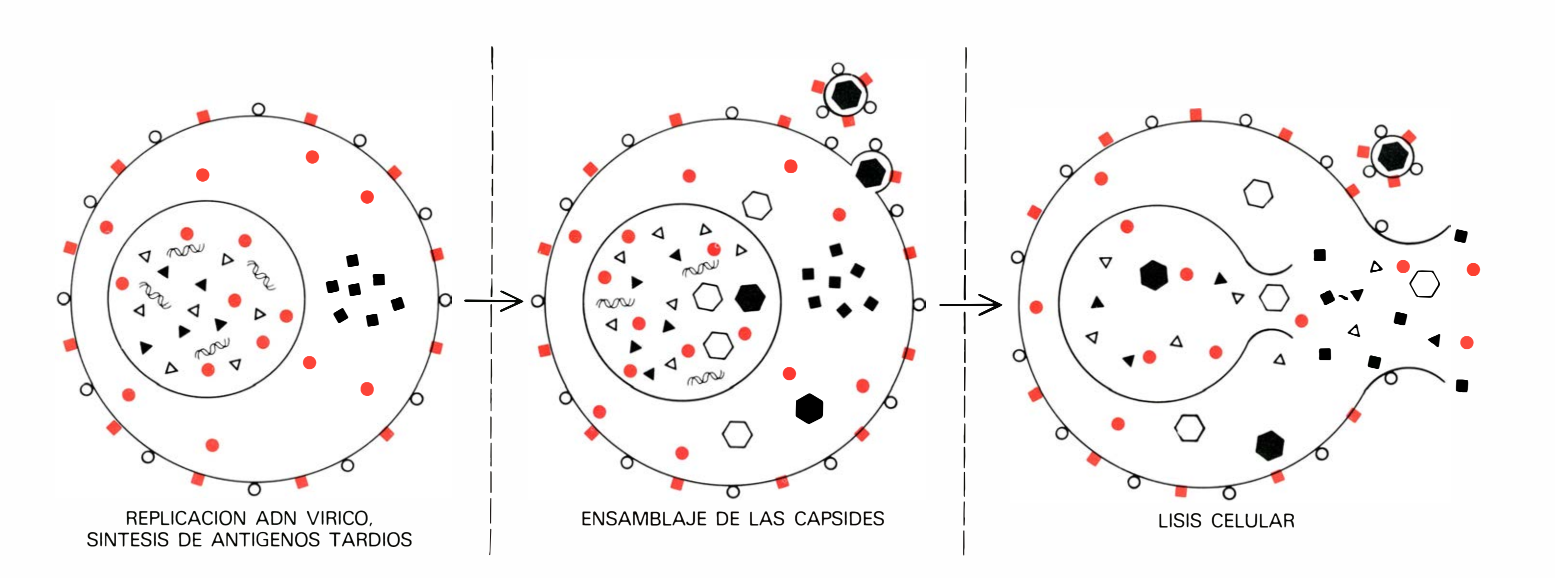
dad de Lund, implica la translocación de un pequeño segmento desde uno de los cromosomas 8 al brazo largo de uno de los cromosomas 14, y de un pequeño segmento del 14 al 8. En otros linfomas se han detectado anomalías del cromosoma 14 similares, aunque no idénticas.

Se desconocen el origen y la función de la aberración cromosómica en el linfoma de Burkitt, pero parece conferir una ventaja biológica a las células afectadas. Por cualquier causa, la aberración puede ser un factor de predisposición o cofactor importante en la génesis del linfoma de Burkitt, ya que los linfocitos obtenidos de enfermos con mononucleosis no presentan cromosomas anormales. El cuadro, sin embargo, no está completo: se han observado translocaciones en muchos, pero no en todos, los linfomas de Burkitt asociados al virus de Epstein-Barr y se han detectado también en células de linfoma de Burkitt no asociadas con el virus. Parece probable que la transformación de linfocitos por el virus de Epstein-Barr constituye sólo la primera fase del proceso de transformación maligna, y que se requieren otros fenómenos para transformar las células.

Experimentos en primates subhumanos han apoyado la hipótesis de la asociación del virus de Epstein-Barr con el linfoma de Burkitt. Thomas Shoppe, D. C. De Chairo e I. George Miller, de la Facultad de Medicina de Yale, inyectaron virus transformantes a tities y otros simios del Nuevo Mundo y observaron varias reacciones. Algunos animales fabricaban anticuerpos contra el virus sin manifestaciones clínicas de enfermedad, otros presentaban una rápida proliferación de linfocitos y otros, por fin, desarrollaban linfomas fatales, en un período de seis a catorce semanas. Las células del linfoma poseían EBNA y los animales fabricaban anticuerpos antiVCA y el complejo antigénico primario. Los linfomas inducidos en simios por el virus de Epstein-Barr, sin embargo, no se ajustan en todos los detalles al linfoma de Burkitt en humanos. Son difusos, en vez de localizados, y se desarrollan en un intervalo de semanas, que no en años, como suele mediar entre la infección vírica inicial en niños africanos y la aparición del linfoma de Burkitt. De todos modos, estos experimentos demostraron claramente que el virus de Epstein-Barr puede inducir tumores malignos.

Los enfermos con linfoma de Burkitt presentan generalmente un título elevado de anticuerpos para VCA y el componente *R* del complejo antigénico primario; en la última fase de la enfermedad pueden desarrollarse anticuerpos anti-componente *D*. Estudios a largo plazo de individuos enfermos han revelado que los pacientes que inicialmente responden a la terapia, si bien conservan o aumentan los títulos de anticuerpos anti-VCA y antígeno *R*, acaban por recaer y aparece por último el tumor de consecuencias fatales, incluso después de períodos que oscilan entre tres y seis años sin rastros patológicos. Por contra, los enfermos que muestran un declive regular en el título de anticuerpos específicos del antígeno *R* después de la cirugía o quimioterapia tienen una gran esperanza de vida a largo plazo sin síntomas patológicos. Los anticuerpos anti-virus de Epstein-Barr son, en consecuencia, buenos indicadores para el pronóstico del linfoma de Burkitt.

Aun cuando todos los niños africanos, casi sin excepción, a la edad de tres años han sido infectados "silenciosamente" por el virus de Epstein-Barr, la incidencia mayor del linfoma de Burkitt



complejo antigénico primario, se detiene la producción de proteínas celulares y la célula muere aunque el proceso se detenga en diversos estadios (*líneas discontinuas*). Muchas células expresan antígenos primarios, pocas los antígenos tardíos y muy pocas células liberan partículas víricas.

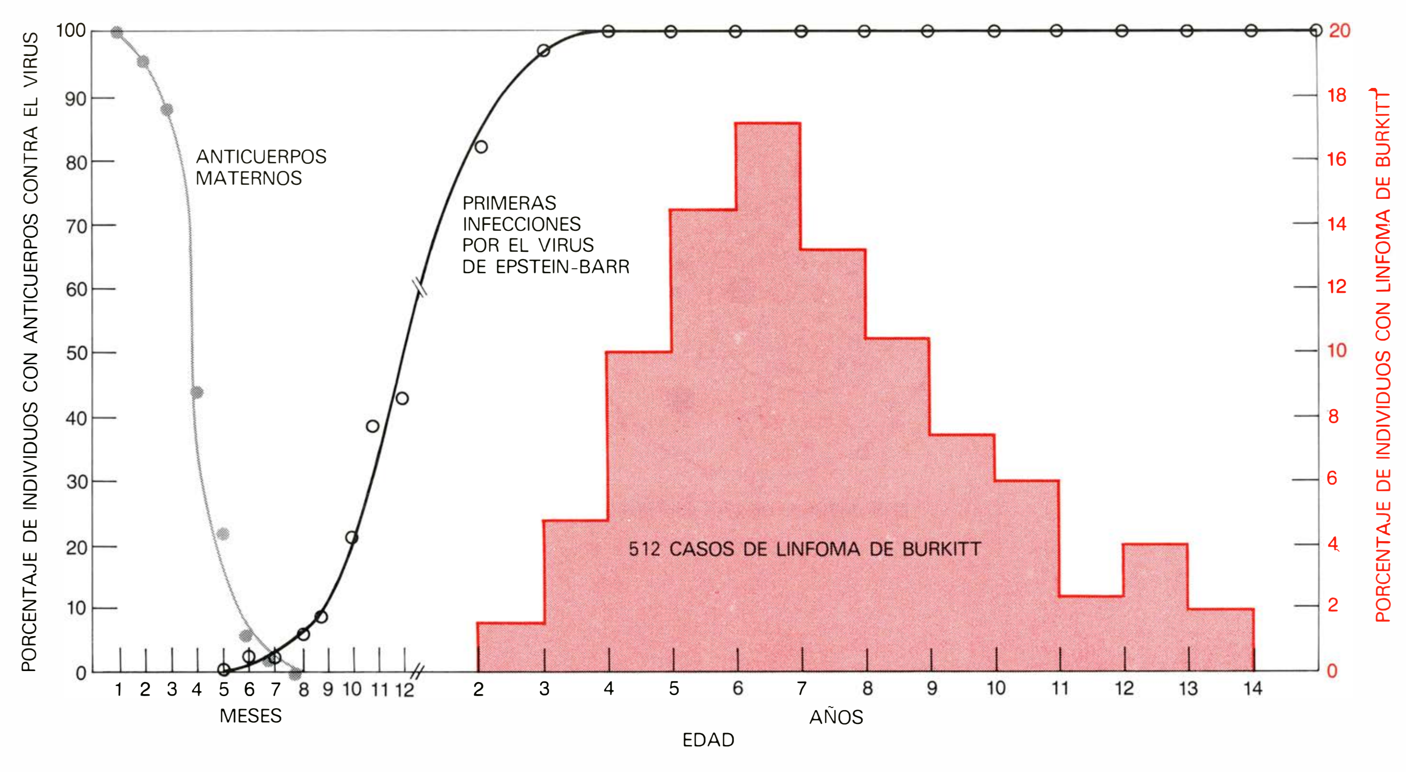
aparece entre los seis y los ocho años. Para investigar acerca de esta aporía desconcertante se organizó en 1972 un estudio prospectivo del linfoma, bajo los auspicios de la International Agency for Research on Cancer in the West Nile District of Uganda, región donde la enfermedad es endémica. Se tomaron muestras de sangre de más de 45.000 niños de edades comprendidas entre uno y ocho años, y se archivaron los sueros para su estudio posterior. Entre 1972 y 1977, 14 de estos niños desarrollaron linfoma de Burkitt. Los sueros obtenidos anteriormente se ensayaron y compararon con sueros procedentes de niños sanos de la misma edad y sexo provenientes de familias vecinas y otros numerosos controles.

Se observó que los niños enfermos habían adquirido anticuerpos contra el virus en un lapso de tiempo entre 7 y 54 meses tras la detección del tumor, y que su título de anticuerpos anti-VCA excedía el título de casi todos los niños sanos control. De cualquier modo, no nos sorprendió que la aparición del tumor no fuera acompañada de un incremento significativo de anticuerpos anti-VCA por encima de los niveles anteriores a la enfermedad. Sin embargo, 10 de los 14 enfermos adquirieron nuevos anticuerpos anti-antígeno *R*. Estos resultados excluyeron la posibilidad de que el linfoma de Burkitt fuera, en los niños africanos, una consecuencia rápida de una extraña y tardía infección por el virus de Epstein-Barr. Llegamos a la conclusión de que un elevado y persistente estado de mantenimiento del virus, puesto de manifiesto por unos niveles altos de anticuerpos, predispone al niño al linfoma de Burkitt, si bien se requieren otros factores para el desarrollo del tumor.

La malaria, altamente endémica en regiones donde el linfoma de Burkitt es corriente, podría constituir uno de esos factores. La infección por parásitos de la malaria puede predisponer a un niño para el linfoma de Burkitt, porque se sabe que la malaria suprime el sistema inmune e intensifica la proliferación de linfocitos. El estudio epidemiológico del efecto de programas de control de la malaria sobre la incidencia del linfoma de Burkitt, que se está desarrollando en la actualidad, habrá de arrojar alguna luz acerca de la relación manifiesta entre estas dos enfermedades. Otro factor importante en la génesis del linfoma podría ser la edad a que ocurre la primera infección por el virus de Epstein-Barr. Infecciones muy precoces podrían transformar linfocitos B inmaduros, incapacitándolos para persistir en el organismo, y dar origen eventualmente a las células del linfoma. En consecuencia, los estudios que habrán de venir sobre estas infecciones prematuras podrían resultar de sumo provecho.

Aunque el linfoma de Burkitt es endémico en varias regiones tanto del Este como del Oeste de Africa, aparece también esporádicamente en otras partes del globo, incluidos los Estados Unidos. Del orden del 80 por ciento de los casos de América del Norte no están asociados con el virus de Epstein-Barr: en las células tumorales no se detectan ni ADN vírico ni células EBNA-positivas, y los enfermos presentan una dotación de anticuerpos comparable a la de los niños normales. El 20 por ciento restante de los casos norteamericanos de linfoma de Burkitt están asociados con el virus de Epstein-Barr y no pueden diferenciarse de los de Africa, excepto en que el nivel de edad en los Estados Unidos se desplaza hacia la adolescencia y la juventud. Recientemente se observó en zonas rurales de Pennsylvania un grupo de jóvenes con linfoma de Burkitt asociado al virus de Epstein-Barr. Los enfermos vivían a unos 48 kilómetros uno del otro, y dos de ellos, que estaban casados, habían compartido la misma casa durante varios años.

Los casos de linfoma de Burkitt que no están asociados con el virus de Epstein-Barr (no sólo el ochenta por ciento de los casos norteamericanos, sino también del orden del dos por ciento de los africanos) presenta claramente un dilema: o el linfoma asociado con el virus es una enfermedad distinta del linfoma que no está asociado con el virus (aunque los tumores sean morfológicamente



INFECCIONES PRECOCES por el virus de Epstein-Barr en niños africanos. Se desarrollan después del descenso de los anticuerpos maternos contra el virus, que son transmitidos al feto a través de la placenta de la madre. En este gráfico se muestra el momento de la infección vírica en relación con este descenso de anticuerpos maternos y la edad de la aparición de quinientos casos de linfoma de Burkitt. Adviértase que la mayor incidencia del tumor se alcanza algunos años después de la infección, lo cual sugiere que la génesis del tumor implica, además del virus, a otros factores distintos.

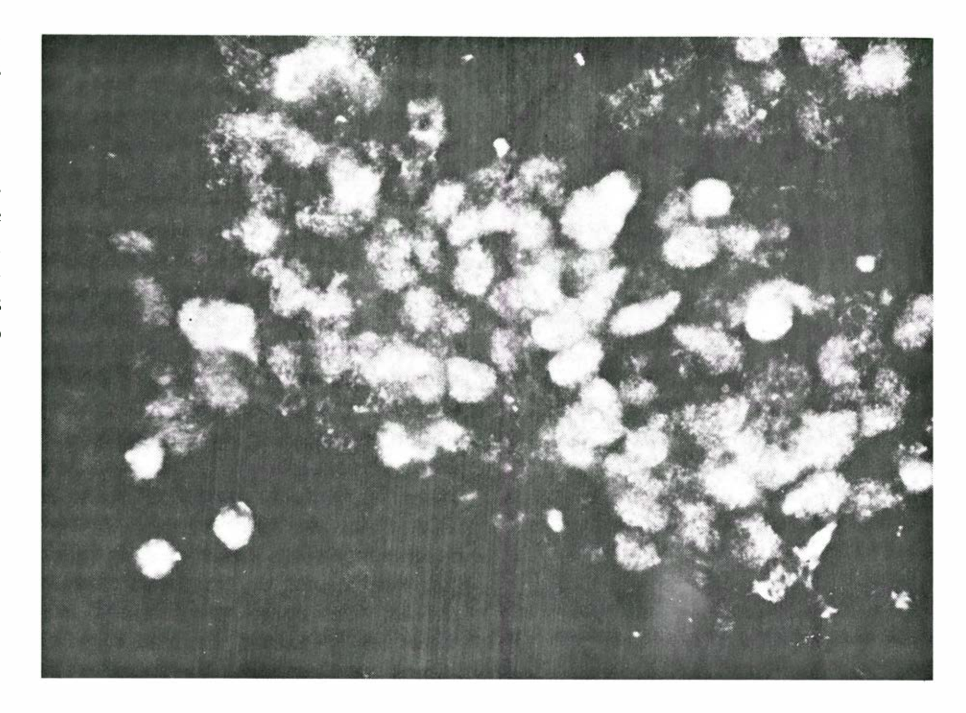
idénticos) o el linfoma de Burkitt tiene por lo menos dos causas, una de ellas ligada al virus de Epstein-Barr. Sea cual sea la respuesta, el factor secundario, al parecer esencial para el desarrollo del tumor, no puede ser la malaria en los casos norteamericanos de linfoma de Burkitt asociado al virus. Sin embargo, no se han reseñado otros factores secundarios ni se ha encontrado ningún virus distinto en los linfomas de Burkitt no asociados al virus de Epstein-Barr.

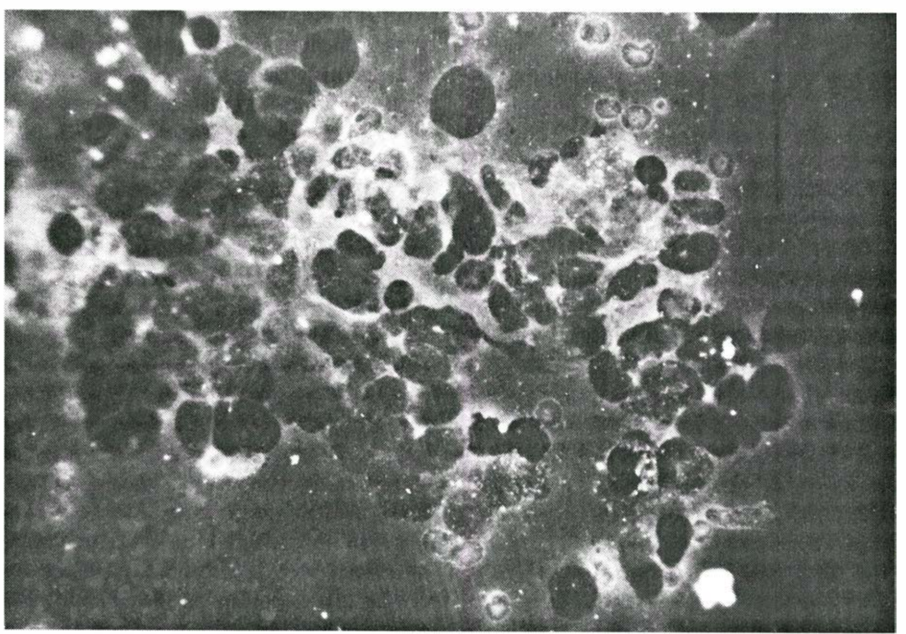
El caso de una asociación entre el virus de Epstein-Barr y el carcinoma nasofaríngeo se basa en criterios similares a los empleados para explicar una asociación entre el virus y el linfoma de Burkitt. Todas las biopsias del carcinoma obtenidas en cualquier parte del mundo han revelado hasta ahora la presencia de ADN vírico que está ausente en otros carcinomas de cabeza y garganta.

Diversos experimentos han demostrado que el ADN vírico está en las células del carcinoma y no en los linfocitos que infiltran el tumor. Por ejemplo, preparaciones de fragmentos de biopsia sobre portaobjetos de cristal pueden poner de manifiesto la presencia de EBNA. No se han detectado, en los tumores, células con infección lítica, quizá porque estas células son destruidas con rapidez por las defensas inmunes del enfermo. Sin embargo, cuando se cultivan las células de carcinoma, puede activarse el genoma vírico latente con fármacos, dando lugar a la fabricación de antígeno *D*, VCA e incluso progenies de virus infecciosos que son capaces de transformar cultivos de linfocitos B. Pero no se han transformado células epiteliales normales, ni se han inducido carcinomas por inyección del virus a primates no humanos. Posiblemente, la razón de estos fracasos estriba en que se usan vías de administración inadecuadas o en que falta un cofactor desconocido.

Los enfermos con carcinoma nasofaríngeo presentan un anticuerpo característico de la respuesta al virus de Epstein-Barr. A diferencia de los enfermos con linfoma de Burkitt, fabrican un título muy elevado de anticuerpos anticomponente *D* del complejo antigénico primario, pero no anti-componente *R*. Otra característica exclusiva del carcinoma es la producción de inmunoglobulina tipo A (IgA). Con frecuencia se presentan niveles elevados de anticuerpos de este tipo específicos para VCA y el antígeno *D*.

En estudios en colaboración con John C. Ho, del Hospital Queen Elizabeth de Hong-Kong, hemos demostrado que ambos títulos de IgG e IgA específicos para VCA y antígeno *D* en enfermos de carcinoma nasofaríngeo aumentan con el incremento de tamaño del tumor, y descienden después de un tratamiento con éxito. En algunos de los enfermos se presentó un descenso inicial en los títulos durante un período de dos a cuatro días, después del cual el curso se invierte y los títulos suben de nuevo. Estos enfermos experimentan una reaparición del tumor o sufren metástasis. El control de los títulos de anticuerpos específicos del virus de Epstein-Barr en enfermos de carcinoma nasofaríngeo proporciona un método para evaluar la eficacia de la terapia y predecir apariciones tumorales inminentes.





PRESENCIA DE UN ANTIGENO VIRICO asociado al virus de Epstein-Barr en una preparación con una muestra procedente de una biopsia de carcinoma nasofaríngeo. Dicha presencia fundamenta la relación entre el virus y el origen de este cáncer humano. Las preparaciones se hicieron con material procedente de biopsias sobre portaobjetos de vidrio. Fueron marcados con anticuerpos fluorescentes anti-antígeno nuclear de Epstein-Barr (*arriba*) y suero privado de anticuerpos específicos (*abajo*). Los núcleos de las células del carcinoma aparecen brillantes en presencia de anticuerpo y oscuros en su ausencia. Esta técnica ha sido desarrollada por los autores en colaboración con Dolly Huang y John Ho.

Estas observaciones sugieren que el virus de Epstein-Barr desempeña un papel de causa en el carcinoma nasofaríngeo, pero, como para el linfoma de Burkitt, quedan muchas cuestiones complicadas sin resolver. Aunque las primeras infecciones también ocurren en los estadios precoces de la vida, en regiones donde el carcinoma es endémico, en China el tumor se desarrolla sólo en

adultos, con un máximo de incidencia alrededor de los 40 años. En Túnez, la distribución de edades es bimodal, con cerca del 20 por ciento de los casos desarrollados antes de los 20 años. El largo intervalo entre la primera infección vírica y el desarrollo del tumor sugiere una vez más la participación de cofactores. A partir de la presentación de varios casos en algunas familias durante tres o cuatro generaciones, en China parece evidente que existe un factor de predisposición. Los chinos afectos de carcinoma presentan un determinado antígeno histocompatible (HLA) con una frecuencia elevada, antígeno que no se observa en enfermos tunecinos. Otro factor concomitante puede ser la ingestión o inhalación de sustancias carcinogénicas, tales como nitrosaminas, presentes en pescados ahumados y en humos de incendios forestales e incienso.

En resumen, el virus de Epstein-Barr es uno de los virus humanos de más amplia difusión. Causa cierta de la mononucleosis infecciosa, todas las pruebas disponibles apuntan a su implicación en la génesis del linfoma de Burkitt y el carcinoma nasofaríngeo. Ambos tumores surgen de la división incontrolada de una única célula maligna transformada, mostrando que el genoma vírico de Epstein-Barr debe haber estado presente en esta primera célula. Parece probable que la verdadera causa de los tumores, cualquiera que sea, tiene por razones inespecíficas preferencia por las células que incorporan el genoma vírico. Tales células, después de todo, representan sólo una ínfima fracción de la población total de linfocitos B a partir de los cuales crece el linfoma de Burkitt. Podría defenderse una situación similar para los precursores de las células del carcinoma. De todos modos es improbable que el virus de Epstein-Barr sea meramente un pasajero inerte de estos tumores. Si es que no es el factor principal, es un inductor.

Parece claro que, para que las células transformadas por el virus devengan malignas, se requieren ciertos sucesos. Estos sucesos pueden ser la acción de carcinogénicos de naturaleza física, química o biológica. Una nutrición deficiente, la acción concomitante de alguna enfermedad tal como la malaria, un descenso en las defensas inmunológicas del huésped o una descompensación en cualquiera de la gran cantidad de funciones orgánicas pueden ser algunos de los factores que contribuyen.

Una vacuna frente al virus de Epstein-Barr, además de conseguir una prevención deseable para la mononucleosis infecciosa, podría aportar datos acerca del posible papel causal del virus en el linfoma de Burkitt y el carcinoma nasofaríngeo, si consiguiera evitar que se presentaran en zonas donde estos tipos de cáncer son endémicos. Sin embargo, el desarrollo de esta vacuna choca con obstáculos quizás insalvables. Por una parte, no puede producirse el virus en grandes cantidades, y hasta ahora sólo puede obtenerse de células con características malignas. Por otra, las vacunas muertas obtenidas de virus herpes han demostrado en general no ser satisfactorias o conferir una protección meramente transitoria. La vacuna viva sería factible si pudiera atenuarse el virus (virus debilitado), pero esta atenuación podría aumentar su capacidad de inducción de cáncer de una forma difícil de predecir.

Hay aún otro problema. En regiones donde el linfoma de Burkitt o el carcinoma nasofaríngeo es endémico, la vacuna debería administrarse durante el período, corto e incierto, en que los antígenos maternos contra el virus de Epstein-Barr (que se transmiten al feto a través de la placenta) han desaparecido de la sangre del niño, y antes de que el niño se infecte de modo natural. Los anticuerpos maternos podrían alterar el efecto de la vacuna, inducir la fabricación de nuevos anticuerpos por el propio sistema inmune del niño. La infección natural por el virus puede producirse poco después de que los anticuerpos maternos hayan desaparecido. Esta cuestión cronológica plantea un problema estratégico que sería difícil de resolver bajo las condiciones de campo imperantes en la zona africana del linfoma de Burkitt. No obstante, deberían hacerse esfuerzos para superar estos obstáculos técnicos, aunque sólo fuera para prevenir la mononucleosis infecciosa.

Esta descripción sirve para ilustrar las dificultades inherentes a suministrar una prueba contundente de la implicación de los virus como causantes de distintos tipos de cáncer humanos. A pesar de que el virus de Epstein-Barr satisface los criterios indirectos que relacionan virus y cáncer humano, quedan incertidumbres acerca del papel exacto del virus en las dos alteraciones malignas. En realidad no puede obtenerse una prueba final que satisfaga a todos. Sin embargo, para los que se inclinan a aceptar las pruebas indirectas, el virus de Epstein-Barr constituye el principal candidato para ser el primer virus canceroso humano conocido.

El parsec central de la galaxia

Las observaciones de radio y en el infrarrojo indican que el centro de nuestra galaxia aloja un objeto ultracompacto, posiblemente un enorme agujero negro, que se halla inmerso en una densa masa de estrellas, gas y polvo en rotación

Thomas R. Geballe

A principios del decenio de 1930, Karl Jansky, joven físico empleado en los Laboratorios de la Compañía Telefónica Bell en Nueva Jersey, construyó el primer radiotelescopio para ayudarse en la identificación de ciertas fuentes de interferencia que podían perturbar las comunicaciones radiotelefónicas transatlánticas. Una importante fuente de descargas que Jansky descubrió fue el "silbido continuo", como él lo llamó, que provenía del cielo y presentaba su máxima intensidad en una dirección muy hacia el sur de su instrumento. A lo largo del tiempo, observó que la fuente de esta interferencia seguía el movimiento de las estrellas fijas y consiguientemente se originaba fuera del sistema solar. La dirección de la máxima intensidad, hacia la constelación de Sagitario, coincidía bien con lo que ya se consideraba por los astrónomos como la posición del centro de nuestra galaxia.

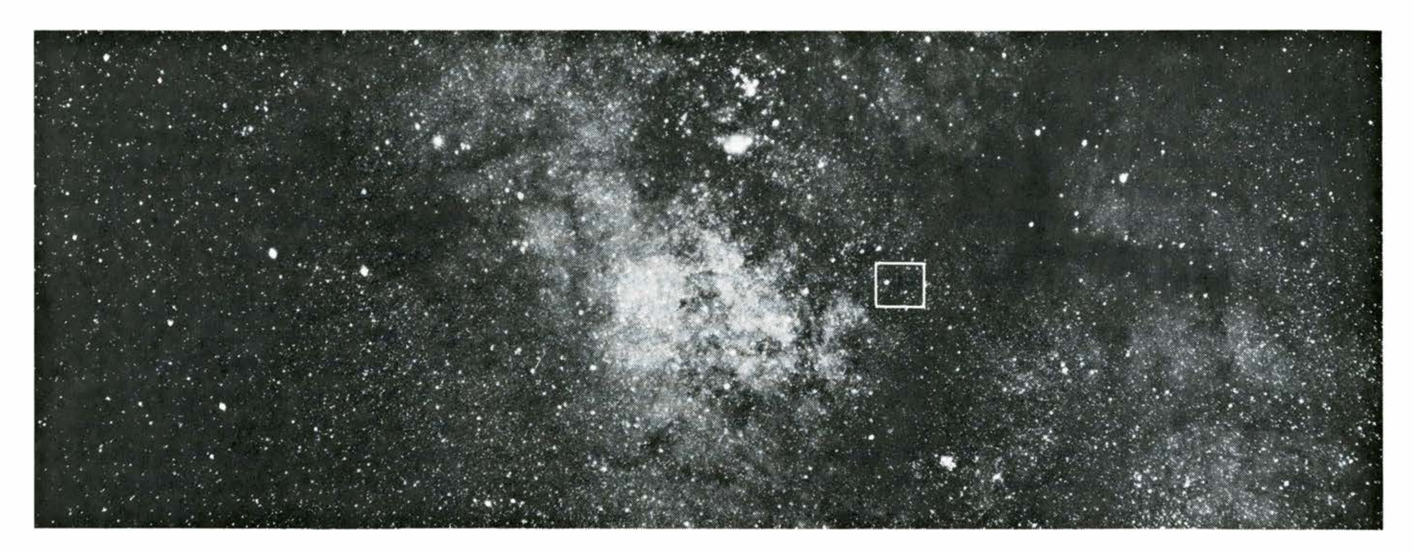
Las observaciones de Jansky fueron no solamente las primeras observaciones radioastronómicas, sino también la primera detección a cualquier longitud de onda del núcleo de la galaxia, que está parcialmente recubierto de nubes de polvo interestelar. Más tarde, aprovechando los avances en electrónica realizados durante la segunda guerra mundial, la radioastronomía ganó la consideración de disciplina científica por derecho propio. En los decenios de 1960 y 1970, la tecnología del infrarrojo, que había quedado retrasada respecto a su contrapartida la de radio, empezó a suministrar detectores sensibles y espectrómetros que se podían usar en combinación con grandes telescopios ópticos. A causa de su capacidad de penetrar las nubes interpuestas de polvo, la radioastronomía y la astronomía de infrarrojo han proporcionado la mayor parte de los conocimientos adquiridos acerca de la región central de la galaxia. En los pocos años últimos, los continuos avances en ambos campos han capacitado a los astrónomos para comenzar a explorar en detalle la región más interna y extraordinariamente densa.

Los núcleos galácticos son densos en masa solamente como consecuencia de la gravitación. Se cree que las galaxias del universo se formaron a partir de heterogeneidades de densidad en el gas expulsado por la "gran explosión" que creó el universo. Con el tiempo, las heterogeneidades se contrajeron por efecto gravitatorio para formar galaxias, y dentro de ellas, otras nubes más pequeñas sufrieron colapsos gravitatorios dando lugar a estrellas y otros objetos. Todos los objetos en la galaxia están también sometidos a una atracción gravitatoria neta hacia su núcleo. Aunque la rotación de la galaxia reduce la caída hacia dentro de la materia, y ciertas explosiones momentáneas (por ejemplo supernovas) pueden expulsar materia hacia el exterior, las fuerzas de rozamiento inevitablemente producen en el centro una densidad másica mucho mayor que la de las regiones externas de la galaxia. Los objetos ultraluminosos conocidos como quasars, observables a enormes distancias, pueden ser un resultado extremo del proceso. El centro, incluso en una galaxia espiral de aspecto normal como la nuestra, hay que esperar que aloje la casi totalidad de la fauna astronómica galáctica: estrellas corrientes, gigantes rojas, enanas blancas, nebulosas gaseosas, nubes de polvo, estrellas de neutrones y quizás incluso un agujero negro, todo ello estrechamente ligado en un compacto enjambre.

Los núcleos galácticos son probablemente la fuente de las enormes expulsiones de materia observadas en algunas galaxias distantes. Parece que la energía necesaria para tales expulsiones solamente puede provenir de los densos núcleos galácticos. Nuestra galaxia también muestra indicaciones de antiguos cataclismos explosivos, aunque relativamente modestos, en las grandes nubes de gas que, según lo observado por los radioastrónomos, se están moviendo hacia el exterior alejándose del núcleo.

Quizá los más sensacionales descubrimientos que se están realizando acerca del núcleo de nuestra galaxia sean los referidos al parsec central, un volumen de solamente 3,26 años-luz de diámetro. A la distancia a que nos encontramos del centro de la galaxia, un parsec subtiende un ángulo de 20 segundos de arco, que es igual al ángulo subtendido en la Tierra por un cráter de tamaño medio en la Luna. Dentro de ese volumen, que encajaría cómodamente entre el Sol y su más próximo vecino estelar, hay millones de estrellas y una gran variedad de otros objetos cuya naturaleza están ahora empezando a comprender los astrónomos. También hay indicaciones de que una cantidad de materia de varios millones de veces la masa del Sol está concentrada en el mismo corazón del parsec central, quizás en la forma de un agujero negro. Aunque el agujero negro no puede observarse directamente, el comportamiento de la materia circundante sugiere su existencia.

He elegido la unidad de un parsec para dar esta información, no porque tenga ninguna virtud inherente sobre alguna otra unidad, tal como el año-luz, sino simplemente porque el parsec es la unidad de distancia habitual del astrónomo. Un parsec es la distancia de la Tierra a un objeto cuya posición, cuando se observa desde la Tierra, parece oscilar a un lado y otro de su posición media en un ángulo de un segundo de arco (1/3600 grados) como consecuencia del paralaje al viajar la Tierra alrededor del Sol. La distancia al centro de la galaxia es de unos 10.000 parsecs, o bien unas veinte veces mayor que la distancia a la bien conocida Gran Nebulosa de Orión. Por otra parte, el centro de la galaxia está unas 70 veces más próximo al sistema solar que el núcleo de cualquier otra galaxia. Por consiguiente, los detalles del núcleo de nuestra galaxia se deberían percibir 70 veces más claramente que los

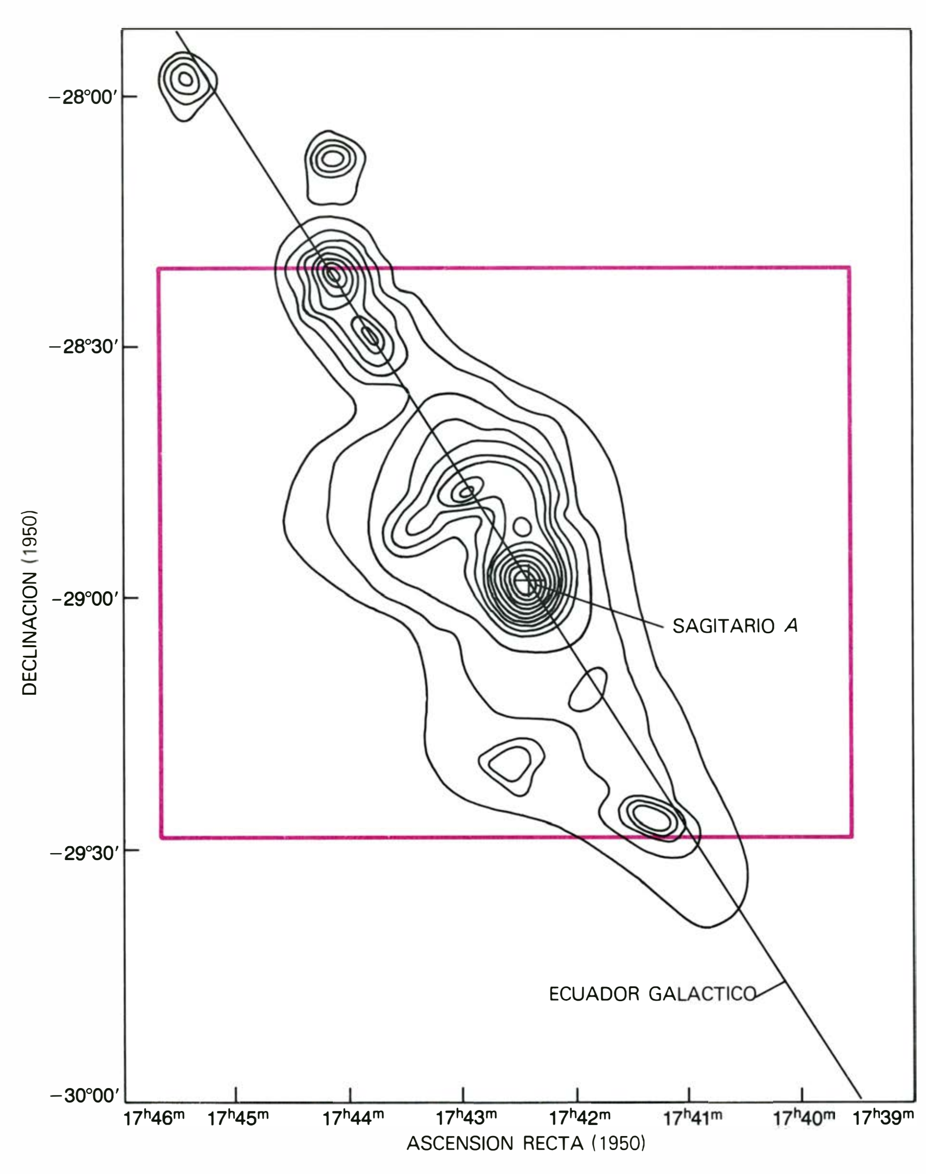


EL CENTRO DE LA GALAXIA queda oculto en la parte Sur de la Vía Láctea en la dirección de la constelación Sagitario. La distancia del sistema solar al centro de la galaxia es de unos 10.000 parsecs, o algo más de 30.000 años-luz. La luz de los millones de estrellas en el núcleo de la galaxia queda completamente oscurecida por el polvo interestelar. Sin embargo, la emisión en las bandas de radio e infrarroja procedente del núcleo galáctico es detectable. El rectángulo, que tiene unas dos veces el diámetro de la luna llena, delinea el área ofrecida en el mapa infrarrojo de abajo.

VISTA INFRARROJA DEL CENTRO GALACTICO, hecha a una longitud de onda de 2,2 micrometros (cuatro veces la longitud de onda de la luz verde) por Eric E. Becklin y Gerry Neugebauer, del Instituto de Tecnología de California y los Observatorios Hale. Su instrumento de observación era el telescopio de un metro en el Observatorio de Las Campañas, en Chile. En esta representación, generada por ordenador en falso color, la intensidad de la radiación aumenta del azul al rojo. La cruz marca el centro de la galaxia. Puede verse que la radiación infrarroja tiende a concentrarse alrededor del ecuador galáctico, indicado por la línea blanca delgada. A 2,2 micrometros, la principal fuente de radiación detectada del centro galáctico son las estrellas rojas y frías. Los distintos puntos brillantes representan gigantes rojas en la parte anterior. El aspecto apedazado es debido al polvo interestelar.

del núcleo de cualquier otra galaxia. Una dificultad, por supuesto, es que cuando se enfoca un telescopio hacia el centro de nuestra galaxia, es imposible distinguir entre los objetos que se hallan en la parte anterior y los de más al fondo. Hay, sin embargo, varios artificios que permiten hacer tal distinción y decidir qué cosas son las que realmente están en el centro galáctico.

Los mayores telescopios y las emulsiones fotográficas más sensibles difícilmente llegan a registrar más que una tenue traza de la luz visible emitida por los millones de estrellas apiñadas en el interior del núcleo de nuestra galaxia. A causa de las nubes de polvo interpuestas, solamente un fotón de luz visible en unos 10^{11} fotones es capaz de sobrevivir al viaje de 30.000 años desde el centro galáctico a la Tierra. Hasta los raros fotones supervivientes quedan empantanados en la radiación de fondo que limita la exposición y que está formada de fotones visibles difundidos en la atmósfera terrestre. En la región infrarroja del espectro, sin embargo, es mucho más elevada la proporción de los fotones que completan su largo viaje. En la región de radio, los fotones procedentes del núcleo galáctico sólo raramente son difundidos y absorbidos. En el otro extremo del espectro, en la región de los rayos X, los fotones también viajan sin impedimento, pero no sobreviven al paso a través de la atmósfera terrestre y por tanto sólo se pueden registrar mediante instrumentos que han sido elevados a gran altura. Sin embargo, nuestra atmósfera es transparente para la mayor parte de las frecuencias de radio y algunas de las del infrarrojo. Como resultado, la mayor parte de lo que se ha averiguado acerca de la región central de la galaxia ha sido suministrado por radiotelescopios y detectores infrarrojos sensibles colocados en los focos de telescopios reflectores.

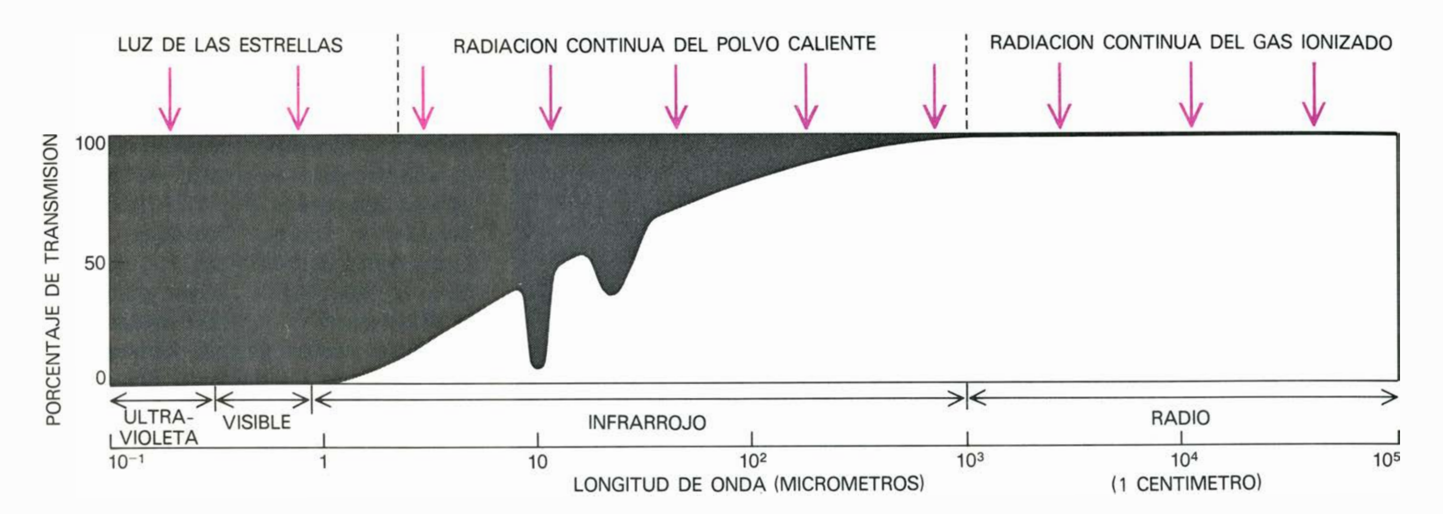


RADIOMAPA DEL CENTRO GALACTICO, hecho a la longitud de onda de 3,75 centímetros por Dennis Downes y Arthur E. Maxwell, de la Universidad de Harvard, y M. L. Meeks, del Instituto de Tecnología de Massachusetts, con el radiotelescopio de 36,6 metros de Westford. La principal fuente de la radioemisión es el gas ionizado. El rectángulo encierra la región representada en la fotografía hecha en longitud de onda del infrarrojo de la página anterior. El gas ionizado y las estrellas están concentrados a lo largo del ecuador galáctico. Presentan su máximo brillo en la radiofuente Sagitario *A*.

Si nuestra galaxia estuviera libre de polvo, sería posible usar un gran telescopio óptico para registrar en una sola placa fotográfica la totalidad del parsec central de la galaxia con una resolución de aproximadamente un segundo de arco. Con un telescopio en un satélite artificial, la resolución se podría aumentar en un factor de 10 a 20. Esta mejora podría bastar justamente para resolver unos 100.000 objetos estelares dentro de un diámetro de 20 segundos de arco, correspondiente al parsec central de la galaxia. Si el parsec central contiene en realidad varios millones de objetos estelares, como parece ahora probable, se agruparían en brillantes racimos, y quizás en una única masa brillante, aun cuando se fotografiasen desde un satélite artificial a alta resolución.

Se puede adquirir una impresión del aspecto que tendría una vista del núcleo galáctico si estuviera libre de polvo mirando fotografías del núcleo de la Gran Nebulosa de Andrómeda, la más próxima galaxia espiral semejante a la nuestra. Por razón de la manera como las dos galaxias están orientadas en el espacio, se puede fotografiar el núcleo de la galaxia de Andrómeda a lo largo de una visual que está en gran parte libre de polvo, el cual se concentra en los planos centrales de los dos sistemas.

El hecho de que el centro de nuestra galaxia es invisible a las longitudes de onda de la luz visible ha constituido sin duda un estímulo para tratar de observarlo a otras longitudes de onda. Para obtener una "fotografía" a las longitudes de onda de radio e infrarrojo, es necesario explorar sucesivas tiras de cielo (por ejemplo, moviendo el telescopio) al mismo tiempo que se registra punto a punto el número de fotones que inciden sobre el detector. De esta manera se puede construir una a modo de imagen de televisión con una resolución que es simplemente el área del cielo que el detector ve en cada instante. El límite de resolución viene impuesto por la apertura efectiva del telescopio: cuanto mayor sea el cociente de la apertura a la longitud de onda, tanto mayor la máxima resolución alcanzable. En la mayoría de los casos, los datos adquiridos de esta manera se convierten en mapas de isolíneas en que las líneas unen puntos



FUENTES DE RADIACION en el núcleo galáctico. Dichas fuentes abarcan estrellas, polvo caliente y las nubes de gas ionizado conocidas como regiones H II. El término H II describe átomos de hidrógeno a los que la radiación ultravioleta procedente de estrellas calientes y próximas ha disociado en protones y electrones. Los electrones y protones emiten radiación en una amplia banda de longitudes de onda (el espectro continuo) al ser acelerados por fuerzas eléctricas, sobre todo las que ejercen entre sí. La curva muestra la proporción de radiación continua que sobrevive al paso a través del polvo interestelar para llegar a los instrumentos en la Tierra. Antes de la radio y la astronomía de infrarrojo, el núcleo era inobservable.

de igual brillo, de la misma manera que las líneas en un mapa topográfico conectan puntos de igual altura. Recientemente, con ayuda de ordenadores, se ha hecho relativamente fácil convertir tales datos en imágenes bidimensionales codificadas en color.

Antes de pasar al parsec central de la galaxia, describiré brevemente lo que se ha aprendido a partir de estudios de radio e infrarrojo acerca de la vasta región nuclear en que se encuentra. Se han construido radiomapas de isolíneas de esta región a muchas longitudes de onda. Todos los mapas revelan una distribución de radiación que es alargada según el ecuador galáctico. La posición del punto de máximo absoluto, conocido por los astrónomos como Sagitario *A*, también es común a todos los mapas, y hace tiempo que se viene sospechando que puede coincidir con el centro galáctico. Es muy probable que Sagitario *A* contribuyera al silbido continuo detectado por Jansky, aunque su instrumento carecía de la resolución necesaria para distinguirlo de otros máximos cercanos.

Las fuentes de emisión que aparecen en los radiomapas del centro galáctico son las nubes de gas interestelar caliente conocidas como regiones H II. Estas regiones, que son corrientes en nuestra galaxia, están formadas por átomos ionizados, es decir, átomos a los que les ha sido arrancado uno o más de sus electrones. Los átomos, principalmente los de hidrógeno, son ionizados por la radiación ultravioleta procedente de las estrellas próximas (u otras fuentes), que son mucho más calientes que el Sol. Los fragmentos atómicos cargados positivamente y los electrones libres, cargados negativamente, de una región H II emiten radioondas en un amplio intervalo de longitudes de onda cuando son acelerados, por ejemplo cuando se mueven en la inmediata proximidad de otras partículas cargadas. En contraste, el gas eléctricamente neutro, tal como el que existe en las regiones H I (que en gran parte está formado por hidrógeno en forma atómica) o en las nubes moleculares (en su mayor parte formadas por hidrógeno en forma molecular), emite poca energía en la banda de radio, excepto en unas pocas longitudes de onda discretas.

La imagen de la vasta región nuclear que se obtiene de las observaciones hechas en el infrarrojo a la longitud de onda de 2,2 micrometros, aproximadamente cuatro veces mayor que las longitudes de onda visibles, es bastante diferente. Mientras la radioemisión proviene del gas interestelar y caliente, la radiación de 2,2 micrometros viene predominantemente de las estrellas. La mayor parte de la radiación de 2,2 micrometros procede de las gigantes rojas, que son estrellas grandes, luminosas y muy avanzadas en su evolución, con temperaturas de superficie tan bajas como 2000 grados Kelvin, en comparación con los 6000 grados K del Sol. Aunque sólo un pequeño porcentaje de las estrellas en el centro galáctico, a lo más, sean quizá gigantes rojas, estos objetos deben estar distribuidos de manera relativamente uniforme entre las estrellas más jóvenes y brillantes, y por tanto deberían suministrar en principio una buena indicación de la distribución de todas las estrellas en su vecindad.

En la parte inferior de la página 37 aparece una imagen en falso color de la vasta región central de la galaxia a 2,2 micrometros. Parte del aspecto apedazado de la imagen se debe a variaciones de la cantidad de polvo interestelar entre el sistema solar y las diferentes porciones del centro galáctico. Los objetos de máximo brillo y apariencia puntual son probablemente gigantes rojas y están, sin duda, en la parte delantera. Esta deducción se confirma por su distribución relativamente uniforme en toda la fotografía.

Aun con las anteriores excepciones, está claro que la emisión de 2,2 micrometros, lo mismo que la radioemisión, está concentrada a lo largo del ecuador galáctico. La dirección de máximo brillo a 2,2 micrometros coincide con un intenso máximo de radio dentro de Sagitario *A*, que se conocen por Sagitario *A* Oeste. Los mapas correspondientes a longitudes de onda mayores en el infrarrojo, como las emitidas por el propio polvo interestelar, también tienen sus máximos cerca de Sagitario *A* Oeste. De aquí que las concentraciones de estrellas, gas caliente y polvo más frío son todas ellas máximas en la misma dirección. Hay poca duda de que el verdadero centro de la galaxia cae en esa dirección.

La masa de estrellas en el centro de la galaxia se puede estimar de las observaciones realizadas a 2,2 micrometros por Eric E. Becklin y Gerry Neugebauer, del Instituto de Tecnología de California y los Observatorios Hale. Becklin y Neugebauer hallaron que la intensidad y distribución de la emisión de 2,2 micrometros procedentes del centro de la galaxia de Andrómeda, designada por M31, es comparable con la emisión procedente del centro de nuestra galaxia. Como el centro de la galaxia M31 se puede fotografiar claramente a longitu-







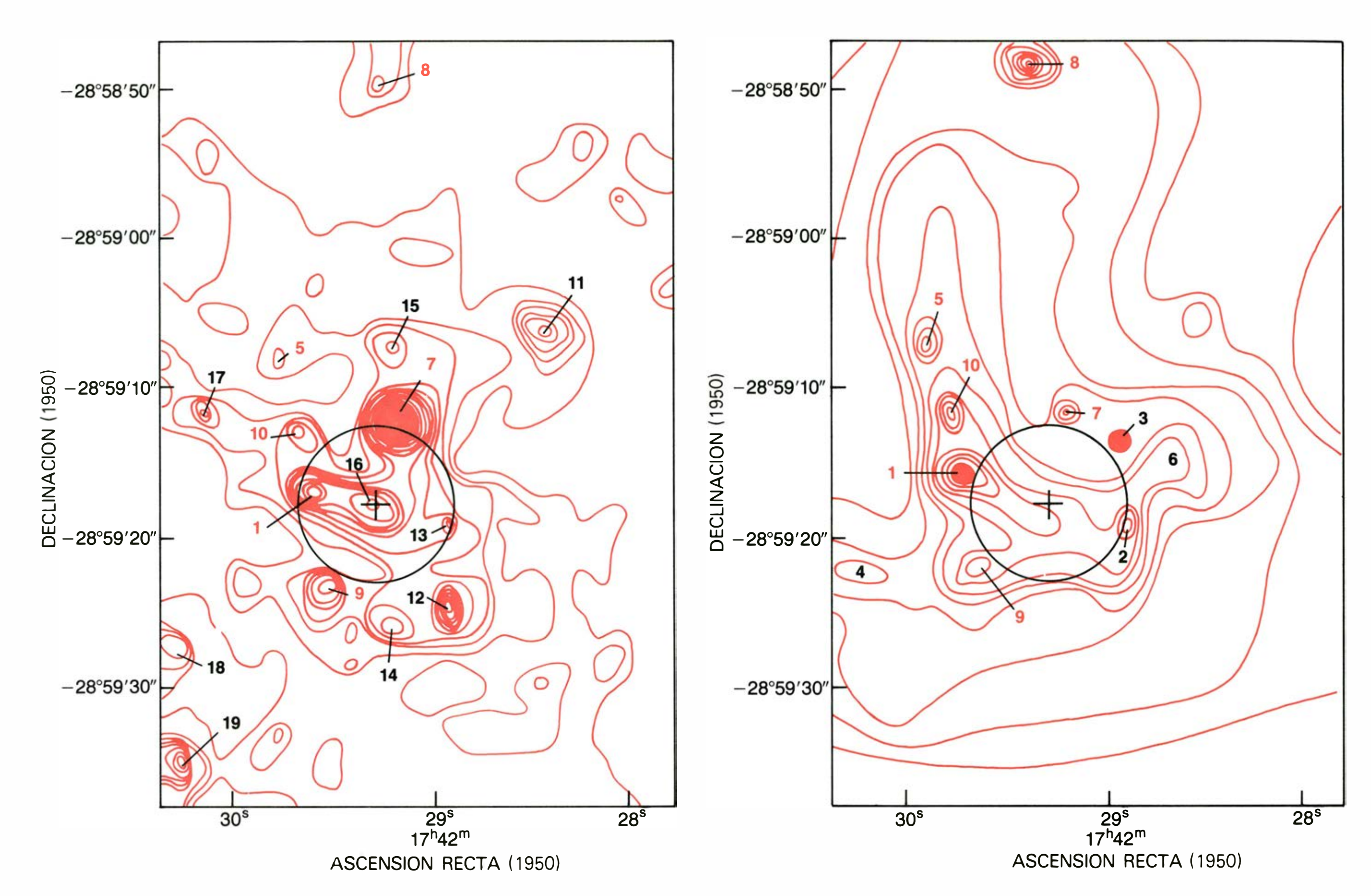
NUCLEO DE UNA GALAXIA ESPIRAL PROXIMA, la Gran Nebulosa de Andrómeda, también conocida como Messier 31 (M31), que puede ser fotografiado a longitudes de onda visibles aun cuando está 70 veces más lejos que el núcleo de nuestra propia galaxia porque hay muy poco polvo interpuesto que atenúe la luz que pasa. En M31, lo mismo que en nuestro sistema espiral, el polvo está concentrado a lo largo del plano del disco galáctico. Las dos galaxias están orientadas en el espacio de tal manera que el núcleo de M31 se ve sin obstrucción. En la fotografía de arriba, que muestra la totalidad de la galaxia, la región central ha sido notablemente sobreexpuesta. La región dentro del rectángulo aparece en la fotografía de la parte baja izquierda, realizada con exposición de un minuto con el telescopio Hale de 5 metros del Monte Palomar. Aunque la mayor parte de las estrellas que casi llenan el área del rectángulo en la fotografía de arriba son ahora invisibles, las miriadas de estrellas muy juntas dentro de los 100 parsecs centrales están todavía sobreexpuestas. La fotografía en la parte baja de la derecha muestra los 20 parsecs centrales del núcleo, que es el área encerrada en el rectángulo en la segunda fotografía. La región borrosa de la parte central de la tercera fotografía, cuyo diámetro mide unos cinco parsecs, representa la luz de millones de estrellas. Los puntos blancos en otras partes no son estrellas, sino granos de la emulsión fotográfica muy aumentados. La fotografía fue tomada con un telescopio de 20 centímetros montado en un globo por E. S. Light, R. E. Danielson y Matin Schwarzschild, de la Universidad de Princeton. Las observaciones infrarrojas sugieren que las estrellas del núcleo de la Gran Nebulosa de Andrómeda y el núcleo de nuestra galaxia están distribuidas de manera análoga.

des de onda visibles, podemos estimar con aceptable exactitud el número de estrellas que deben estar allí representadas. Suponiendo que la relación entre radiación infrarroja y radiación visible es la misma para el centro de nuestra galaxia que para el de la M31, Robert H. Sanders y Thomas Lowinger, de la Universidad de Columbia, y J. H. Oort, de la Universidad de Leiden, calcularon la distribución de la masa estelar en el centro de nuestra galaxia. Estiman que una esfera de un diámetro igual a un parsec centrada en el brillante máximo central de la emisión de 2,2 micrometros contiene una masa en estrellas equivalente a unos dos millones de soles.

El cúmulo globular de estrellas más densamente apretado fuera del núcleo galáctico tiene unas 10.000 estrellas dentro del parsec central. A distancia de un parsec del Sol no hay otras estrellas. Suponiendo que cada estrella del parsec central sea igual que nuestro Sol, la estima anterior conduce al resultado de que la distancia media entre estrellas en el parsec central es de aproximadamente una semana-luz, es decir, unas 100 veces la distancia entre el Sol y el planeta Plutón. Si la Tierra estuviera en el parsec central, los millones de estrellas próximas bañarían el planeta con una cantidad de luz igual a la de varios cientos de lunas llenas. El que la vida pudiera existir en el interior del parsec central, lleno de radiación, es otro asunto.

Para observar una región tan pequeña y distante como el parsec central de la galaxia es deseable trabajar con instrumentos que tengan la máxima resolución angular posible. Los mayores radiotelescopios que operan en longitudes de onda cortas tienen una resolución limitada a aproximadamente un minuto de arco, resolución suficientemente baja para difuminar por completo cualquier estructura espacial en el parsec central. Usando dos o más radiotelescopios separados por una gran distancia y empleando la técnica interferométrica, sin embargo, la resolución se puede aumentar substancialmente. Por desgracia, la posición austral del centro galáctico lo ha hecho casi inaccesible a la interferometría con los grandes radiotelescopios que hasta hace poco sólo existían en latitudes más altas.

Sin embargo, la situación presente habrá de mejorar muy pronto en cuanto se complete la instalación de radio conocida como Distribución Muy Grande (Very Large Array) en Socorro, Nuevo México. Este instrumento permitirá, por fin, explorar el parsec central con una resolución angular comparable con la



MAPAS INFRARROJOS del centro galáctico construidos a dos longitudes de onda en el infrarrojo, 2,2 micrometros (*izquierda*) y 10,6 micrometros (*derecha*), que ponen de manifiesto características distintas. La circunferencia de cada mapa representa lo que se cree ser el parsec central de nuestra galaxia. Sólo seis de las más prominentes fuentes infrarrojas aparecen en ambos mapas: 1, 5, 7, 9 y 10. Algunas de las fuentes que brillan a 2,2 micrometros son, casi con certeza, gigantes rojas, estrellas que son al mismo tiempo frías y extremadamente luminosas. Las IRS (del inglés *infrared source*, fuente de radiación infrarroja) 7, IRS 11 e IRS 12 son gigantes rojas. Exceptuando IRS 3 e IRS 7, la mayor parte de las fuentes en el mapa de 10,6 micrometros se cree que son nubes calientes de polvo asociadas con regiones H II. La cruz en los dos mapas coincide con el brillante objeto radioemisor dentro de Sagitario *A* Oeste, que probablemente rodea un agujero negro. El mapa a 2,2 micrometros fue obtenido por Eric E. Becklin y Gerry Neugebauer con el telescopio de 5 metros. El mapa adjunto de 10,6 micrometros es una composición de varios mapas obtenidos con diversos telescopios, y cuya realización se debe a G. H. Riecke, de la Universidad de Arizona, y C. M. Telesco y D. A. Harper, de la Universidad de Chicago.

que ahora se puede obtener en las longitudes de onda del infrarrojo.

La alta resolución posibilitada en longitudes de onda de radio por la interferometría se ha aplicado ya, de una manera limitada, para estudiar Sagitario *A* Oeste. En 1974, Bruce Nalick, a la sazón en el Observatorio Lick, y Robert L. Brown, del Observatorio Nacional de Radioastronomía en Green Bank, Virginia Occidental, descubrieron por radiointerferometría que Sagitario *A* Oeste tiene cerca de su centro un objeto radioemisor que es, a un tiempo, excepcionalmente brillante y excepcionalmente pequeño. Desde entonces, se ha demostrado que esta compacta radiofuente tiene un núcleo central cuyo diámetro no es mayor de 0,001 segundos de arco, o escasamente diez veces la distancia de la Tierra al Sol (suponiendo que el objeto está en realidad en el centro de la galaxia).

Aunque el compacto radioobjeto presenta alguna semejanza con ciertas clases de estrellas radioemisoras, también reúne algunas de las propiedades que la materia debería presentar cuando ha sido atraída al interior de un agujero negro. Se cree que puede formarse un agujero negro cuando una estrella de gran masa que ha agotado su combustible nuclear es ya incapaz de generar la presión interna necesaria para equilibrar la fuerza atractiva de la gravedad. Bajo tales circunstancias, el interior de la estrella se contraerá hasta hacerse tan pequeño y denso que las leyes conocidas de la física no pueden aplicársele. La fuerza gravitatoria es tan grande en la proximidad de un objeto así contraído, que hay una región a su alrededor de la cual esencialmente nada, ni siquiera la luz, puede escapar; esta región es un agujero negro.

La materia que se encuentre sometida al intenso campo gravitatorio que reina en las proximidades de un agujero negro será acelerada hasta adquirir grandes velocidades y se calentará hasta temperaturas muy altas como efecto de los choques. Como resultado, un agujero negro debe estar rodeado de una pequeña pero intensa radiofuente en forma de disco, que presenta una apariencia semejante a la de la radiofuente compacta. En el centro galáctico, donde el gas y el polvo son abundantes, debe esperarse que un agujero negro devore la materia a un ritmo comparativamente elevado, creciendo con rapidez en masa y en influencia gravitatoria. La posibilidad de que Sagitario *A* Oeste aloje un enorme agujero negro en el centro de la galaxia ha entrado sin duda en el pensamiento de muchos astrónomos. Conforme vayamos acanzando en el artículo iré examinando las indicaciones recientes acerca de esta posibilidad.

El análisis detallado de los pocos parsecs centrales de la galaxia ha resultado mucho más fácil en longitudes de onda en el infrarrojo que en la banda de la radio. El estudio ha sido llevado a cabo a varias longitudes de onda en el infrarrojo a lo largo de la última media docena de años principalmente por G. H. Rieke y Frank J. Low, de la Universidad de Arizona, y por Becklin y Neugebauer. A cada longitud de onda han hallado

una brillante concentración de objetos dentro del diámetro de un parsec coincidente con Sagitario *A* Oeste. A diferentes longitudes de onda, sin embargo, la luminosidad relativa y la distribución de los objetos son sorprendentemente diferentes.

Las diferencias quedan quizá más claramente puestas de manifiesto por las observaciones a 2,2 micrometros y 10,6 micrometros [*véase la ilustración de la página anterior*]. A cada longitud de onda se pueden observar muchos objetos distintos, pero sólo seis de ellos aparecen simultáneamente a ambas longitudes de onda. En otras palabras, los objetos que son prominentes en una longitud de onda no lo son en la otra. Por ejemplo, el objeto conocido como IRS 1 (fuente infrarroja 1), que es el más brillante a 10 micrometros, ostenta sólo un brillo moderado a 2,2 micrometros. IRS 7, con gran diferencia el objeto más brillante a 2,2 micrometros, apenas si se detecta a 10 micrometros. Parece que los objetos observados a distintas longitudes de onda pertenecen a distintas clases.

¿Qué clase de objetos son? Todos ellos son en el infrarrojo más brillantes de lo que aparecerían la gran mayoría de estrellas si las observásemos a una distancia de 10.000 parsecs. Como he mencionado, puesto que las gigantes rojas son particularmente brillantes a 2,2 micrometros, es natural preguntarse si alguna de las fuentes a 2,2 micrometros en el centro galáctico son gigantes rojas. Hay una manera de contestar la pregunta. En las atmósferas de las estrellas más frías es posible la formación de moléculas. Una especie que se forma en abundancia es el monóxido de carbono, que absorbe radiación comenzando a 2,3 micrometros y extendiéndose a longitudes de onda mayores. La absorción tiene lugar más fuertemente en las gigantes rojas luminosas. En 1975, observaciones hechas por grupos en la Universidad de Arizona y en el Cal Tech revelaron vaguadas de absorción a 2,3 micrometros en el espectro de tres de las fuentes en el centro galáctico (IRS 7, IRS 11 e IRS 12), demostrando que cada fuente consta al menos de una gigante roja luminosa. Por otra parte, una de las fuentes que es brillante a 2,2 micrometros, IRS 16, muestra poca o ninguna absorción de monóxido de carbono. La identificación de IRS 16 es todavía problemática. Su posición parece coincidir con la compacta radiofuente de Sagitario *A* Oeste, de modo que éste es de particular interés.

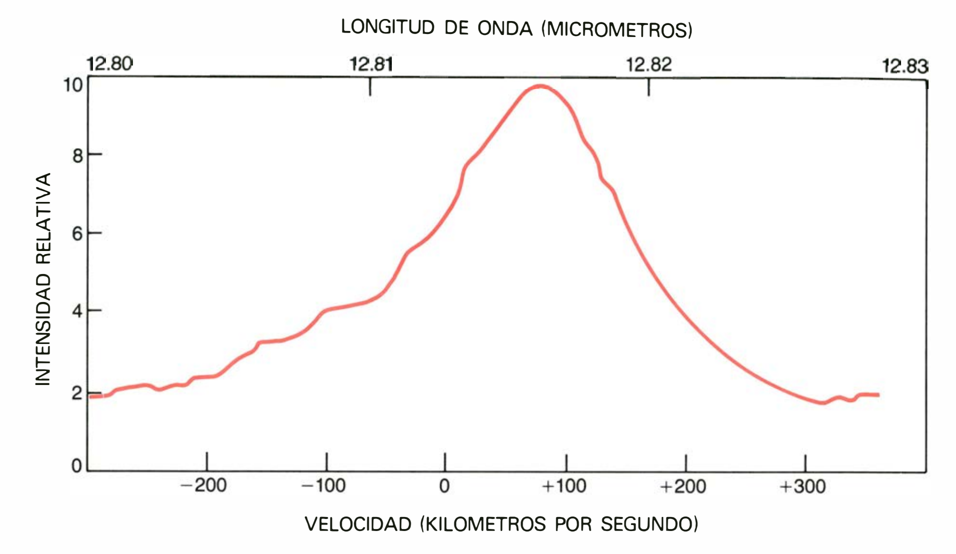
ESPECTRO INFRARROJO DE CINCO FUENTES próximas al centro galáctico; muestra que tres de ellas, IRS 7, IRS 11 e IRS 12, llevan la "marca" de las gigantes rojas: una caída en la intensidad a longitudes de onda superiores a 2,3 micrometros. Tales estrellas tienen una atmósfera rica en monóxido de carbono, que absorbe fuertemente la radiación a 2,3 micrometros y más allá. IRS 1, que es una región H II, e IRS 16, en Sagitario *A* Oeste, no muestran tal absorción. Los espectros fueron obtenidos con el telescopio de 5 metros por Becklin, Neugebauer, Beckwith, Matthews y Wynn-Williams.

Un segundo objeto que no muestra absorción de monóxido de carbono, IRS 1, pertenece a la clase de objetos que son mucho más brillantes a 10 micrometros que a 2,2. Muy poca de la emisión de radiación a 10 micrometros procedente de objetos de esta clase puede ser radiación estelar, porque incluso las estrellas más frías son mucho más débiles a 10 micrometros que a 2,2. Los astrónomos que trabajan en el infrarojo han encontrado ahora muchas fuentes desperdigadas alrededor de la galaxia que son desusadamente brillantes a 10 micrometros. Casi en todos los casos, la fuente de la radiación se ha identificado como polvo caliente. Las partículas de polvo absorben radiación visible y ultravioleta de las estrellas próximas u otros objetos luminosos y reemiten la energía absorbida en longitudes de onda del infrarrojo. Algo semejante ocurre probablemente en el núcleo galáctico, y se puede suponer que los máximos en el mapa a 10 micrometros representan nubes de polvo en o cerca del parsec central.

El gas interestelar de nuestra galaxia es unas 100 veces más abundante en masa que el polvo interestelar, y su presencia siempre queda indicada por la presencia del polvo. Por tanto, parece razonable suponer que el gas radioemisor de Sagitario *A* Oeste está asociado a nubes de polvo y que está calentado por la misma fuente o fuentes que calientan aquellas. Si éste es el caso, sería de esperar que los mapas de alta resolución de Sagitario *A* Oeste revelasen una estructura similar a la que se ve en los mapas a 10 micrometros. Había indicaciones en 1974, deducidas de estudios radiointerferométricos, de que Sagitario *A* Oeste es realmente una reunión de nubes más compactas de gas ionizado. En aquella época, sin embargo, era difícil determinar si algunos de los máximos de emisión en el infrarrojo y los radiomapas estaban o no asociados en el espacio.

Hasta este punto, la historia cubre aproximadamente el estado de conocimientos acerca del núcleo galáctico hacia 1975. Se sabía que dentro del parsec central había una concentración masiva de estrellas cuya luz visible es ocultada

por el polvo. Dentro de unos pocos parsecs del centro se habían encontrado fuentes emisoras en el infrarrojo de diversos tipos. De las fuentes que eran brillantes a 2,2 micrometros, se había demostrado que algunas eran gigantes rojas frías y luminosas; muchas de las otras fuentes podrían también alojar estrellas frías, pero éstas todavía no habían sido identificadas. La mayoría de los objetos que eran brillantes a 10 micrometros se parecían a regiones H II compactas, pero aunque se sabía que había gas ionizado en el centro galáctico, se desconocía si el gas estaba físicamente asociado con alguna de esas fuentes infrarrojas. Se había hallado que una intrigante fuente de radioemisión, que podría ser un objeto supermasivo en el mismo centro de la galaxia, coincidía con una peculiar y brillante fuente que emite intensamente en 2,2 micrometros. Por último, no se sabía nada en absoluto acerca del movimiento de esos objetos; la materia en el parsec central podría haber estado contrayéndose, dilatándose, girando o moviéndose sin orden.



EMISION DEL NEON ionizado con una carga, Ne II, presente en las nubes gaseosas en el parsec central de la galaxia, estudiado por el autor y sus colegas con un espectómetro infrarrojo unido al telescopio de 1,5 metros en el Observatorio Interamericano de Cerro Tololo, en Chile. El examen de la línea de emisión del Ne II cerca de 12,8 micrometros revela que la línea se ha ensanchado apreciablemente como consecuencia del movimiento de las nubes gaseosas, de las que el neón es un componente en pequeña cantidad. El movimiento hacia el observador (llamado negativo) corre la línea hacia frecuencias más altas o longitudes de onda más cortas, mientras que el movimiento alejándose del observador (llamado positivo) produce el efecto opuesto. Las velocidades observadas del neón ionizado oscilan entre aproximadamente −200 a + 200 kilómetros por segundo a lo largo de la visual, proporcionando una indicación de la existencia de intensas fuerzas gravitatorias cerca del centro de la galaxia. La intensidad de la radiación continua de fondo en la vecindad de 12,8 micrometros, procedente del polvo caliente en el núcleo, se deduce a partir del nivel de radiación a cada lado de la línea de neón ensanchada.

Las estrellas, gas y polvo en el centro galáctico emiten radiación cuya variación de intensidad respecto a la longitud de onda es bastante gradual; la radiación se llama consiguientemente fondo continuo. Otra fuente de información potencialmente rica es la radiación de líneas: la radiación emitida y absorbida a frecuencias específicas y características por los átomos y las moléculas. El espectro de los objetos astronómicos generalmente consta de ambas cosas, fondo continuo liso y líneas. Cada línea está asociada a un nivel específico de energía de un átomo o molécula y tiene una cierta intensidad y anchura y una longitud de onda central.

La intensidad de la línea es una medida de la temperatura y de la abundancia de la substancia emisora. Si la fuente se está moviendo hacia el observador, la línea se corre hacia longitudes de onda más cortas; si la fuente se está alejando, la línea se corre hacia mayores longitudes de onda. Tales corrimientos Doppler informan al astrónomo acerca de la velocidad radial o velocidad de una fuente a lo largo de la visual. Las velocidades hacia el observador ("corrimientos azules") se llaman negativas; las que se alejan del observador ("corrimientos rojos") se llaman positivas. Cuando los átomos o moléculas fuentes se mueven a velocidades variables, los corrimientos varían; la línea se difumina y se dice que tiene una anchura. Por tanto, la anchura de una línea puede indicar el margen de velocidades radiales de la fuente.

En 1973, David Aitken, Bárbara Jones y James Penman, del University College London, fueron los primeros en detectar una línea de emisión infrarroja que se originaba en el centro galáctico. La longitud de onda de la línea, en la vecindad de 12,8 micrometros, identificaba su fuente como neón ionizado una vez, o Ne II: neón al que le falta uno de sus 10 electrones. Con anterioridad se había observado que la misma línea del neón era una de las líneas de emisión infrarroja más intensas en varias regiones ionizadas de otras partes de la galaxia. Estas regiones también se hacen presentes por emisión continua de radiación en la región de radio del espectro.

Usando la intensidad conocida de emisión de radio procedente de Sagitario *A* Oeste y suponiendo que la abundancia de neón en el gas ionizado de Sagitario *A* Oeste es normal, Aitken, Jones y Penman calcularon la intensidad que debería tener la línea de neón. La intensidad observada de la línea resultó ligeramente mayor que la prevista, lo que significaba que el gas en el centro galáctico es más rico en neón que el gas que se observa en los demás lugares. La predicción era sin embargo suficientemente aproximada para que pudiera deducirse que el mismo gas ionizado produce tanto la línea infrarroja de neón como el fondo continuo de radioemisión. De aquí que el trabajo de Londres demostrara que la línea de 12,8 micrometros de Ne II, que representa sólo una pequeña fracción (aproximadamente 0,01 por cien) del gas ionizado, podría servir de trazador para el estudio del movimiento y distribución del gas ionizado de Sagitario *A* Oeste, que se suponía estar en el mismo centro de la galaxia.

El grupo de Londres trabajó con un instrumento de infrarrojo que permitía la entrada de un haz circular cuyo diámetro correspondía aproximadamente a un parsec en el centro galáctico. La resolución de su espectrómetro era demasiado baja para determinar la anchura de la línea de neón y si mostraba o no corrimiento Doppler. En el Departamento de Física de la Universidad de California en Berkeley, donde yo era entonces estudiante de los últimos cursos de licenciatura, una sucesión de estudiantes que trabajaron con Charles H. Townes habían estado desarrollando espectrómetros de infrarrojo capaces de alta resolución espectral. En 1975, Eric Wollman, John Lacy y yo, junto con Townes y David Rank, de la Universidad de California en Santa Cruz, medimos la forma de la línea de Ne II con el telescopio de 1,5 metros del Observatorio Interamericano de Cerro Tololo en Chile (donde Sagitario pasa directamente

por encima). Lo que encontramos puso en claro que Sagitario *A* Oeste no es una región galáctica H II típica. La línea de neón tenía una anchura de varios cientos de kilómetros por segundo, mientras que lo normal para una región H II son velocidades en la vecindad de 30 kilómetros por segundo.

Por la misma época, Thomas Pauls, Peter G. Mezger y Edward Churchwell, del Instituto Max Planck de Radioastronomía en Bonn, realizaron radioobservaciones de las líneas de emisión generadas por la recombinación de los núcleos de hidrógeno y electrones que también mostraron que el gas ionizado en Sagitario *A* Oeste estaba en rápido movimiento. No se había determinado todavía si éste era organizado o turbulento ni cómo estaba distribuido en el parsec central el neón y el resto del gas ionizado.

La respuesta a estas y otras preguntas requería observaciones de la línea de 12.8 micrometros a una escala espacial más fina. En 1975 y 1976, en el Observatorio Lick, observamos varias posiciones dentro del parsec central y hallamos que tanto la intensidad de la línea como la velocidad radial variaban con la posición. En 1977, Lacy completó un espectrómetro nuevo y más sensible que hizo posible observar la línea Ne II todavía con mayor detalle. En 1977 y 1978, Lacy, Townes, Fred Baas (visitante de Leiden) y yo, trabajando con este instrumento en el nuevo telescopio Du Pont de 2,5 metros en el Observatorio de Las Campañas en Chile, obtuvimos una cobertura completa de la línea Ne II dentro del parsec central y mas allá de él. Trabajando con diámetros de apertura correspondientes a fracciones pequeñas, del orden de 0,2 parsec, en el centro galáctico, obtuvimos espectros en los vértices de una red rectangular y construimos mapas de isopletas de la intensidad de la línea a varias velocidades diferentes.

Nuestros datos muestran que hay una gran diversidad tanto en la intensidad de la línea de neón como en la velocidad de éste observada en el parsec central. En unas pocas direcciones, la intensidad de la línea muestra máximos para más de una velocidad. Sagitario *A* Oeste no se puede considerar como una región uniforme de gas ionizado, sino que más bien debe estar compuesta de un cierto número de fuentes independientes más pequeñas que se mueven unas respecto a las otras. Como he indicado, gran parte del fondo continuo de radiación infrarroja a 2,2 y 10 micrometros procede también de grupos de pequeñas fuentes. Era importante determinar si la distribución del gas ionizado estaba relacionada con alguno de los mapas de 2,2 o 10 micrometros.

Para investigar este problema, anotamos la posición de los máximos de intensidad del Ne II a ciertas velocidades. Cuando inscribimos las posiciones de las componentes de velocidad sobre los mapas del fondo continuo infrarrojo, encontramos que las máximas intensidades coincidían con muchas de las fuentes que eran brillantes a 10 micrometros. Por ejemplo, IRS 1 está en la misma posición que un máximo de la línea de neón a la velocidad cero. Cerca de IRS 2 hay poco neón ionizado a velocidad nula, pero hay un agudo máximo en el neón ionizado que se está acercando al observador a una velocidad de 280 kilómetros por segundo, velocidad no observada en ninguna otra parte del parsec central. Al menos hay otras cinco fuentes infrarrojas, IRS 4, IRS 5, IRS 6, IRS 9 e IRS 10 que coinciden con componentes de velocidad localizadas correspondientes a la línea Ne II.

Las coincidencias espaciales de gas caliente inizado y polvo caliente son indicaciones convincentes de que la mayor parte de las fuentes brillantes del mapa del fondo a 10 micrometros constituyen regiones H II compactas dentro de Sagitario *A* Oeste. De la intensidad de las líneas de neón podemos estimar que la materia ionizada en cada una de esas regiones equivale aproximadamente a una masa solar. La masa del polvo asociado, que es la responsable de la emisión de fondo a 10 micrometros, es mucho menor. El diámetro de una región compacta H II es típicamente de un cuarto de parsec. Como hay aproximadamente dos millones de estrellas en el parsec central, es probable que cada región compacta H II aloje hasta 100.000 estrellas.

Sólo dos de las fuentes del mapa de 10 micrometros no son, decididamente, regiones H II. Una de ellas, la IRS 7, se sabe ya que es una gigante roja. La otra, IRS 3, es más misteriosa. La IRS 3 no puede ser una gigante roja porque no es brillante a las longitudes de onda más cortas dentro del infrarrojo. Tampoco puede ser una región H II porque hay en ella una clara disminución de la intensidad de la línea Ne II en la vecindad de esta fuente. Ello sugiere que IRS 3 enmascara parte del gas ionizado. Quizá IRS 3 sea una nube fría y opaca de polvo y gas, bien expulsada recientemente por una estrella o bien en proceso de colapso gravitatorio para convertirse en estrella, delante del parsec central.

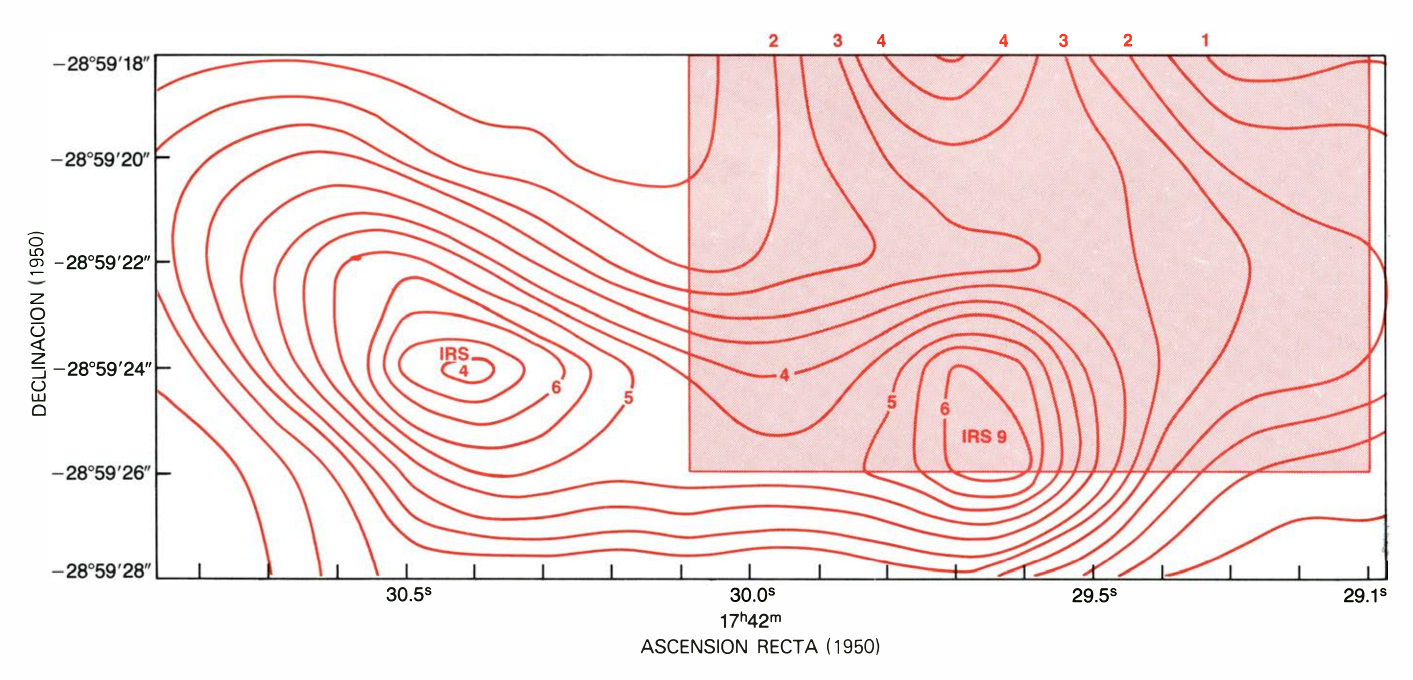
Cada una de las regiones compactas H II del parsec central se está moviendo a una velocidad cuya componente radial viene dada por el corrimiento Doppler de su línea de neón asociada. La mayoría de las regiones H II están claramente separadas en espacio o en velocidad, y deben constituir, por tanto, objetos independientes. Sin embargo, el estudio de todos los espectros de neón indica que hay un movimiento colectivo del parsec central. Esto se ve claramente observando que una línea que vaya aproximadamente de norte a sur a través de la parte media de la región separa el neón ionizado que está predominantemente acercándose del que está predominantemente alejándose. La más sencilla interpretación de esta división de velocidades es una rotación: la línea divisoria correspondería entonces al eje de rotación.

Aunque hasta ahora sólo se han encontrado indicaciones de rotación en el gas ionizado, parece razonable que el resto de la materia del parsec central (incluyendo las estrellas) se esté moviendo de una manera muy parecida a como lo hace el gas ionizado. Basándose en los corrimientos Doppler de la línea de neón observados a unas décimas de parsec, del eje aparente, se puede estimar que el tiempo que tarda el gas, polvo y estrellas dentro del parsec central en realizar un giro completo alrededor del centro galáctico es de unos 10.000 años. (El sol tarda unos 200 millones de años.) La orientación del eje es peculiar: es casi perpendicular al eje de rotación del disco galáctico principal. Esto sugiere que la región nuclear es una estructura separada, distinta de sus alrededores. No se sabe todavía a qué distancia del centro se extiende el peculiar eje de rotación.

Los movimientos del gas ionizado en el parsec central, tanto si son parte de una rotación colectiva como si no, sufren la influencia de la fuerza gravitatoria de la masa que se encuentra dentro de tal volumen y puede utilizarse para calcular dicha masa. Como el resto de la galaxia está distribuido más o menos simétricamente alrededor del centro, con la mayor parte de su masa a gran distancia de él, casi no ejerce fuerza gravitatoria neta sobre el parsec central. La masa total del parsec central, estimada a partir de las velocidades de las líneas de neón, vale entre cinco y ocho millones de veces la masa del Sol. De ese total, según parece deducirse del estudio de las observaciones a 2,2 micrometros, sólo dos millones de masas solares se encuentran en el estado de estrellas de todas las clases. La masa total de polvo y gas dentro de la región es despreciable en

−28°59′14″
−28°59′17″
−28°59′20″
−28°59′23″
−28°59′26″
DECLINACION (1950)
IRS 10
IRS 1
RADIOFUENTE COMPACTA
IRS 6
IRS 2
IRS 9
30.10^{s} 29.85^{s} 29.60^{s} 29.35^{s} 29.10^{s} 28.85^{s} 28.60^{s}
$17^{h}42^{m}$
ASCENSION RECTA (1950)
−400 0 +400

ESTUDIO DETALLADO de la línea de neón. Nos ilustra la complejidad del movimiento dentro del parsec central de la galaxia. El retículo de espectros se ha ordenado de acuerdo con su posición en el cielo, e incluye un poco más del parsec central. En el vértice inferior derecho, la circunferencia da el tamaño del haz del espectómetro y la escala da la velocidad del gas en kilómetros por segundo. Para cada espectro, la posición del origen coincide con una velocidad nula a lo largo de la visual, con intensidad de emisión cero y con la posición a la que se apuntó el telescopio. Varias de las componentes de velocidad de la línea de neón que se han localizado coinciden con fuentes que son brillantes a 10 micrometros (IRS 1, IRS 2, IRS 6, IRS 9 e IRS 10), indicando que el gas ionizado está asociado físicamente con nubes de polvo calientes. La línea de trazos inclinada separa regiones de neón ionizado que están predominantemente acercándose (*derecha*) y alejándose (*izquierda*). Una separación tan bien definida sugiere que la materia en el parsec central está girando en el sentido indicado por la flecha. La región dentro del rectángulo en color vuelve a aparecer en la figura de abajo.



INTENSIDAD DE LINEA DE NEON en la parte sureste del parsec central. Demuestra la asociación del neón ionizado con dos fuentes específicas IRS 4 e IRS 9, cuya radiación continua se origina en nubes de polvo calientes. Las curvas muestran la intensidad del neón con velocidades de unos +125 kilómetros por segundo. Los máximos coinciden con las dos fuentes infrarrojas. Los espectros tomados en la posición de las fuentes indican que IRS 4 está en realidad alejándose a 110 kilómetros por segundo e IRS 9 a 140 kilómetros por segundo. Las observaciones correspondientes a las ilustraciones de esta página fueron realizadas por el autor y sus colegas con el nuevo telescopio Du Pont de 2,5 metros en el Observatorio de Las Campañas, en Chile.

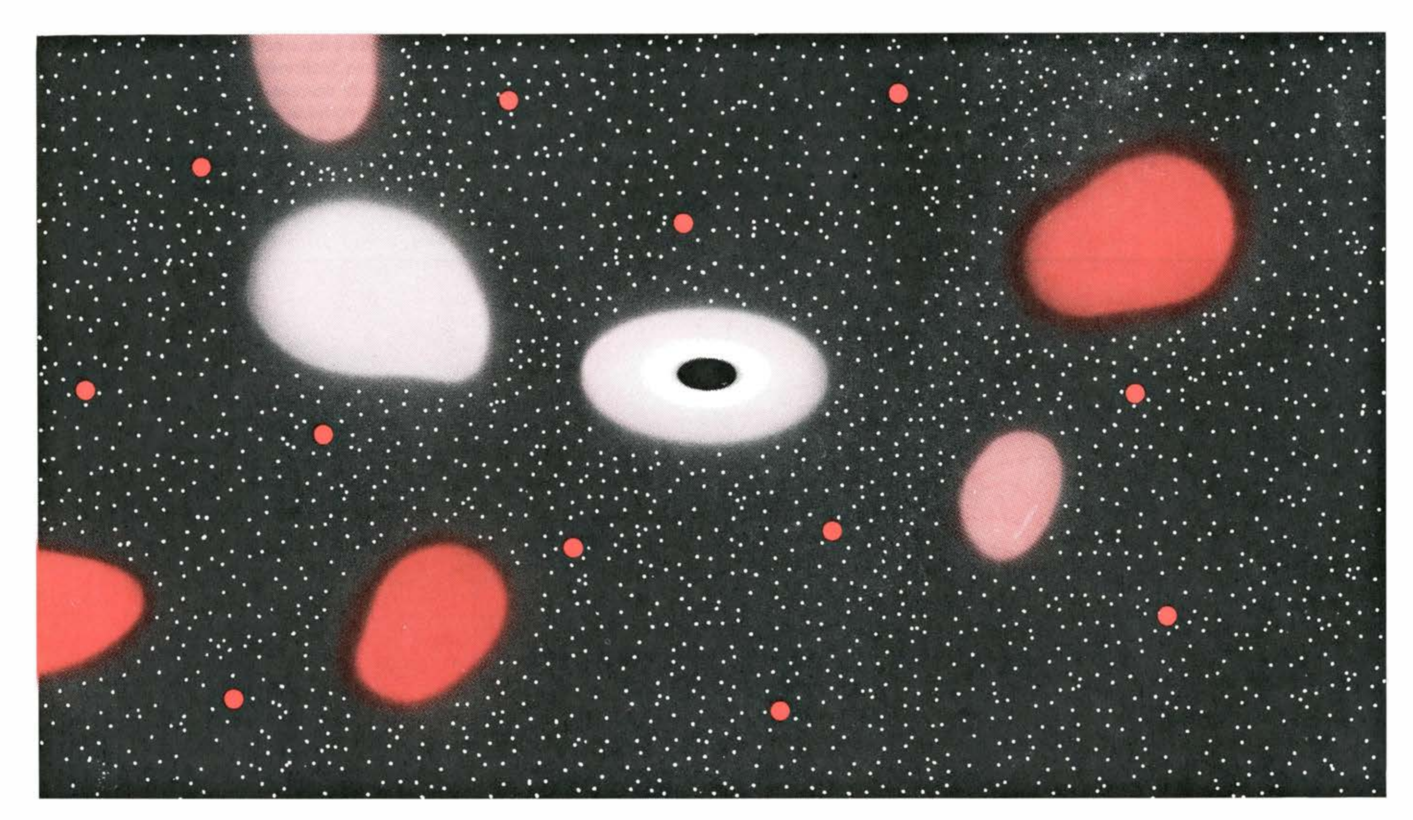
comparación: probablemente no más de unas 10 masas solares.

Resulta fascinante, pues, especular sobre si la diferencia de entre tres millones a seis millones de masas solares que hay entre la masa estelar aparente y la masa total es significativa. ¿Representa masa desconocida todavía por considerar? ¿O son los cálculos inexactos? Hay una observación que puede ser pertinente. Las mediciones de la línea infrarroja del neón y de las radiolíneas generadas por la recombinación de núcleos de hidrógeno y electrones muestran que dentro de unos pocos parsecs centrales las velocidades del gas tienden a aumentar hacia el centro. Luis Rodriguez y J. Chaisson, del Centro de Astrofísica del Harvard College Observatory y del Observatorio Astrofísico Smithsoniano, y también algunos otros, han hecho notar que tal aumento no puede ser debido a la fuerza gravitatoria del cúmulo central estelar a menos que éste tenga un núcleo desusadamente denso o incluya un objeto central compacto, masivo y presumiblemente de naturaleza no estelar.

Al considerar los modelos teóricos para explicar las observaciones del parsec central, supongamos (al menos inicialmente) que el núcleo no se está observando en una época particular. Las observaciones han demostrado que el parsec central contiene un cúmulo de estrellas extremadamente denso junto con una cantidad apreciable de polvo y gas ionizado, que están distribuidas en forma irregular y se encuentran en rápido movimiento. Un modelo satisfactorio debe ser capaz de explicar no sólo la presencia de tales propiedades, sino también su presencia continuada.

A causa de la atracción gravitatoria hacia el centro de la galaxia, no es sorprendente hallar tal centro poblado por un denso cúmulo de estrellas. Las esporádicas nuevas aportaciones al cúmulo son probablemente estrellas ordinarias que habían girado en torno al centro galáctico a una distancia más grande y que han perdido momento cinético, tal vez como consecuencia de un pase próximo a otras estrellas. Inversamente y por medio del mismo proceso, estrellas que se hallan dentro del núcleo pueden ocasionalmente adquirir velocidad suficiente para escapar de él. Consiguientemente, la población estelar del parsec central es probable que permanezca próxima a un régimen estacionario o al menos cambie muy lentamente.

En contraste, las densas concentraciones de polvo y gas ionizado tienen normalmente vida corta. En las regiones exteriores de la galaxia hay dos clases de tales objetos: las regiones H II que se hallan en la vecindad de estrellas calientes y luminosas recientemente formadas, y las nebulosas planetarias, que también son regiones H II pero cuya ionización se debe a ciertas estrellas viejas que son ordinariamente más calientes, aunque menos luminosas. De hecho, se necesitaría la radiación ultravioleta procedente de unas 100 del tipo de estrellas que ionizan las nebulosas planetarias para dar cuenta de la cantidad de gas ionizado que se observa en el parsec central, pero como hemos visto, sólo se observan allí unas 10 nubes ionizadas. Por esta razón se puede objetar la identificación de las nubes como nebulosas planetarias. Por otra parte, para producir ese efecto bastaría con unas 10 de las estrellas luminosas y de reciente formación del tipo que ionizan las regiones H II normales. Otras fuentes típicas de polvo y gas, tales como el "viento" de partículas emitido por las estrellas o la caída de materia hacia el interior procedente de las regiones circundantes, no produciría la distribución de materia estelar extre-



VISTA ESQUEMATICA DEL NUCLEO GALACTICO, de un diámetro de unos tres parsecs, que postula la existencia de un agujero negro de gran masa (dibujado mucho más grande que a escala) en el mismo centro. Las manchas luminosas de color representan nubes de gas y polvo interestelar, creadas a razón de una cada 1000 años, o así, por choques estelares dentro del denso cúmulo central de estrellas. Algo del gas y polvo, atraído por acción gravitatoria hacia el agujero negro, forma un disco de acumulación caliente y giratorio antes de ser devorado y desaparecer. La intensa radiación ultravioleta procedente del disco de acumulación ioniza el gas en las nubes circundantes, que emiten entonces las líneas infrarrojas y la radiación continua en frecuencia de radio que se han detectado en la dirección del parsec central. El polvo en las nubes se calienta también por radiación del disco de acumulación y las estrellas próximas, y emite radiación infrarroja continua. Una fracción importante de la radiación infrarroja de onda corta proviene de unas pocas gigantes rojas (topos de color) dentro del denso cúmulo central. Los puntos blancos simbolizan estrellas corrientes.

madamente poco uniforme que se deduce de las observaciones del fondo continuo infrarrojo y de las líneas de neón.

Sin embargo, en una región tan densamente poblada de estrellas como es el parsec central debe considerarse un mecanismo adicional de producción de nubes densas: los choques estelares. Un choque disruptivo, es decir, en el cual dos estrellas llegan realmente a establecer contacto, arrancaría materia de una de ellas o de las dos, creando una nube turbulenta e inicialmente densa que se dispersa en el medio circundante y más enrarecido. Cálculos realizados por Townes, Lacy y David Hollenbach indican que, dentro del parsec central, los choques en los cuales al menos uno de los participantes es una gigante roja son una fuente importante de materia interestelar. Las nubes que tales choques crearían tienen tamaños, velocidades, movimientos internos y períodos de vida que concuerdan con las propiedades de las nubes que realmente se observan. Además, la frecuencia estimada de tales choques puede ser la adecuada para producir nubes en el número apropiado.

La pregunta siguiente es: ¿cuál es la fuente de radiación ionizante necesaria para convertir las nubes del parsec central en regiones H II? He mencionado las estrellas calientes y luminosas y las nebulosas planetarias como posibles agentes ionizantes. Las estrellas calientes y luminosas permanecen en esa fase durante sólo unos pocos millones de años; por tanto, no pueden haberse trasladado muy lejos del punto donde ahora se observan, que es el propio núcleo galáctico. Sin embargo, se cree que la formación de estrellas es un proceso delicadamente equilibrado que lleva consigo la contracción de grandes nubes de gas y polvo a lo largo de un período de aproximadamente un millón de años. En el turbulento medio circundante del núcleo galáctico, en el que los períodos orbitales pueden ser tan pequeños como 10.000 años, quizá no sea posible que grandes volúmenes de gas y polvo se contraigan para formar estrellas antes de ser completamente dispersados.

La limitación más importante sobre el tipo de fuente de ionización en el parsec central parece ser el uniformemente bajo nivel de ionización en la región. Las regiones H II típicas en las otras partes de la galaxia contienen generalmente, además de neón ionizado con una carga, especies atómicas con varias cargas, es decir, a las que faltan dos o más electrones. Para crear los estados más fuertemente ionizados, se necesitan fotones ultravioleta de alta energía, procedentes de estrellas extremadamente calientes.

En el curso de nuestras observaciones de Ne II, también buscamos intensamente en el parsec central líneas espectrales de argón doblemente ionizado (Ar III) y azufre triplemente ionizado (S IV) en sus longitudes de onda conocidas en el infrarrojo. Esas líneas son frecuentemente muy fuertes en las regiones H II fuera del núcleo galáctico. Los resultados en varias posiciones en el parsec central fueron negativos, con la excepción de una línea muy débil de Ar III detectada en la dirección de IRS 1. Recientemente, Steven P. Willner, Ray Russell y Richard Puetter, de la Universidad de California en San Diego, Thomas Soifer, del Cal Tech, y Paul Harvey, de la Universidad de Arizona, detectaron una fuerte línea infrarroja del Ar II (argón ionizado con una sola carga) procedente del parsec central. Su resultado implica una moderada superabundancia de argón en el parsec central y, combinándolo con nuestro resultado, pone de manifiesto que el estado de ionización en el parsec central es desusadamente bajo.

Las estrellas centrales de las nebulosas planetarias, que son generalmente muy calientes, parecen quedar descartadas como fuente de radiación ionizante porque la mayor parte de ellas dan lugar a altos estados de ionización. A menos que las estrellas más masivas y calientes tengan algún impedimento para formarse en el centro galáctico, o a menos que las más recientemente formadas se hayan enfriado de manera importante, sería de esperar que las estrellas ionizantes presentasen un amplio intervalo de temperaturas. Tal intervalo de temperaturas conduciría a variaciones en el estado de ionización del gas en el interior del parsec central. La ausencia de Ar III y S IV en todo éste parece implicar que o bien todas las estrellas ionizantes son casi idénticas, o bien hay una fuente ionizante en el parsec central que es extremadamente luminosa pero de temperatura relativamente baja. Un objeto de este último tipo no puede ser una estrella ordinaria.

Tales consideraciones conducen a un modelo bastante exótico del parsec central: un modelo dominado por un agujero negro de gran masa en el centro. La materia que está siendo devorada por un agujero negro se cree que se acumula primero alrededor de él en un disco giratorio; se admite que bajo condiciones adecuadas debe radiar fuertemente en el ultravioleta. La uniformidad del estado de ionización del parsec central po-

dría explicarse por la presencia de tal objeto. Un candidato natural para el agujero negro es el compacto radioobjeto de diámetro inferior a 0,001 segundos de arco en Sagitario *A* Oeste.

Las propiedades que debe esperarse tengan los agujeros negros de gran masa han sido calculadas por muchos astrofísicos teóricos. El espectro y la intensidad de la radiación ionizante procedente del disco depende sólo de la masa y de la velocidad de aumento de materia del agujero negro. Townes, Lacy y Hollenbach han demostrado que se puede obtener concordancia cuantitativa entre la luminosidad observada y el estado de ionización de la región H II en el parsec central de la galaxia postulando un agujero negro que tuviera una masa de unos cuantos millones de veces la del Sol e incrementara su materia al moderado ritmo de aproximadamente 10^{-5} masas solares por año. La masa necesaria de materia puede ser de sobras suministrada por caída hacia el interior, vientos estelares, nebulosas planetarias y restos de choques estelares en el parsec central. Aun concediendo que parte de la materia sea arrastrada hacia fuera por la intensa presión de radiación del disco de almacenamiento, de los quizá dos millones de estrellas y de otros objetos luminosos en el parsec central, parece que quedaría una cantidad de materia ampliamente suficiente para alimentar el agujero negro al ritmo requerido.

Quizás es importante el hecho de que la intensidad de radiación infrarroja emitida por el peculiar objeto IRS 16, en el que la compacta radiofuente parece estar situada, es aproximadamente la misma que la que se emitiría en la vecindad de un agujero negro que acumulase materia a la velocidad arriba indicada. Por último, la existencia de un agujero negro con una masa de varios millones de soles está de acuerdo con la diferencia entre la masa total estimada en el parsec central y la masa de las estrellas en el mismo volumen. Un agujero negro de tal masa cabe en un volumen del tamaño de la radiofuente compacta.

Las indicaciones de la presencia de un agujero negro masivo son razonables, aunque indirectas. Comprobaciones más directas podrían provenir de futuras observaciones en las bandas de radio e infrarrojo de la pequeña región alrededor de las radiofuentes compactas. Por ejemplo, el intenso campo gravitatorio en la vecindad de un agujero negro de tal masa debería dar a las líneas espectrales que se originasen en la fuente una anchura Doppler de 1000 kilómetros por segundo o más. Cuando nuestro espectrómetro infrarrojo, cuya anchura de haz es de 0,2 parsec, se centró en la radiofuente compacta, que es una posible posición del agujero negro, la línea de neón no pareció tener esa anchura. Sin embargo, puede haber existido una débil y ancha componente que nuestro instrumento no haya conseguido detectar. Los cambios importantes en la velocidad de acumulación de materia en el agujero negro, los cuales pueden producirse en el plazo de unos pocos años, deberían dar lugar a cambios observables en el estado de ionización del parsec central. De aquí que sea importante la vigilancia continua de todas las líneas de emisión infrarroja.

De nuevo podría uno preguntarse: ¿Estamos realmente observando el parsec central de la galaxia? En mi opinión, la respuesta es resueltamente afirmativa. En primer lugar, estamos apuntando nuestros telescopios en la dirección del parsec central; ello lo indican todas las observaciones en el infrarrojo y radio, quizá con la máxima claridad las de 2,2 micrometros de la distribución estelar. En segundo lugar, la rica distribución de fuentes de radiación infrarroja observada en esta dirección tiene pocas contrapartidas, si es que hay alguna, en otras partes de la galaxia. Las altas velocidades positivas y negativas obtenidas a partir de la línea de neón son únicas, y las enormes densidades de masa que se deducen de las observaciones de 2,2 micrometros y de las líneas de neón no se han observado en ninguna otra parte.

Es posible que el parsec central sea la fuente de fenómenos que se observan mucho más allá del núcleo galáctico. Los bruscos y grandes incrementos en la velocidad de acumulación del agujero negro postulado, las supernovas o explosiones de formación de estrellas cerca del centro galáctico pueden ser la causa de las nubes moleculares que se mueven hacia fuera y han sido observadas por los radioastrónomos. Tales explosiones expulsarían materia del núcleo galáctico en forma de nubes de alta velocidad. Con el tiempo, la materia chocaría con las nubes moleculares observadas y las aceleraría; tales nubes, a su vez, podrían ser los restos de explosiones anteriores.

Aunque el núcleo de nuestra propia galaxia es una potente fuente de energía, es muy débil comparado con muchos otros núcleos galácticos. Quizás esa es una de las razones por la que estamos aquí y en condiciones de estudiar el núcleo galáctico. En el presente puede estar pasando por una fase de calma entre épocas más explosivas. El estudio continuado del núcleo galáctico conducirá a una mejor comprensión del mismo.

Ciencia y sociedad

Agujeros en la historia

Se diría que los agujeros negros, esos supuestos fantasmas de estrellas masivas colapsadas, a partir de las cuales la no materia, luz o señal de cualquier tipo puede escapar, son como un hijo cuya paternidad exclusiva pertenece a la cosmología moderna. La verdad es que semejante creación se remonta a la década de 1770, escribe Simon Schaffer, de la Universidad de Cambridge, en un ensayo reciente publicado en *Journal of the History of Astronomy.*

John Michell, astrónomo del siglo XVIII, que fue párroco de Thornill, Yorkshire, analizó la atracción de la luz por parte de la gravedad en una comunicación aparecida en 1784 en *Philosophical Transactions of the Royal Society of London.* En un artículo que llevaba por título "Sobre los medios para descubrir la distancia, la magnitud y otros parámetros de las estrellas fijas, a partir de la diminución de su luz, siempre que ocurra una tal pérdida en cualquiera de ellos, y con tal que los datos vengan suministrados por los observatorios si nuestro propósito los necesitara", Michell sostenía que si la luz estaba sujeta a la gravedad, la velocidad de la luz que emitía una estrella quedará frenada por la atracción gravitacional de ésta. Consiguientemente, podríamos determinar la masa de una estrella midiendo la deceleración de la luz estelar. De acuerdo con los cálculos de Michell, la velocidad de huida del Sol es 497 veces más pequeña que la velocidad de la luz. La velocidad de escape es proporcional al radio, y así afirmaba que una estrella dotada de la densidad solar y un radio 497 veces mayor atraparía toda la luz que emitiera.

En 1796 el matemático y físico francés Pierre Simon de Laplace se expresaba en el mismo sentido, aunque, al parecer, no tenía conocimiento de la obra de Michell. En su *Exposition du système du monde* Laplace llegó a la conclusión de que toda estrella que tuviera la densidad del Sol y un radio 250 veces mayor podría resultar invisible por cuanto la velocidad de escape en superficie sería mayor que la velocidad de la luz. Los análisis de Michell y Laplace sobre los agujeros negros se basaban, ni que decir tiene, en la mecánica de Newton, en tanto que los estudios contemporáneos parten de la teoría general de la relatividad y la mecánica cuántica. Media pues un abismo entre ambas explicaciones. Michell y Laplace no pasaron de la conjetura, sin fundamento teórico u observacional, sobre la posibilidad de que la luz tuviera propiedades de partícula y sufriera por tanto el influjo de la gravedad. Se sabe hoy con certeza que tal es la verdad. Las leyes de la física moderna exigen que la luz tenga propiedades de partícula, y la prueba experimental ha revelado de qué naturaleza son. También es conocido ya que el mecanismo de operación de un agujero negro no sería la deceleración de la velocidad sino el contenido de las partículas de luz, los fotones, por la curvatura del espaciotiempo.

Umbral más bajo

Como bien saben los cuidadores de perros, sonidos que sobrepasan la capacidad auditiva del hombre son percibidos con nitidez por dichos animales. El límite superior de la capacidad auditiva humana ronda los 20.000 hertz (ciclos por segundo), en 40.000 se halla la de los perros y en 80.000 la del ratón y murciélago. El límite inferior o umbral auditivo se sitúa, en humanos y en situaciones normales, en los 125 hertz. ¿Hay animales que pueden oír por debajo de ese límite inferior?

La respuesta es afirmativa si hemos de creer en los resultados de Rickye y Henry Heffner, de la Universidad de Kansas en Parsons, y Ned Stichman, del Instituto Zoológico Ralph Mitchell de Independence. Según parece, todos los animales con cabeza superior a la del hombre poseen esa capacidad. Esos tres investigadores han sometido a prueba la gama auditiva del elefante hindú de siete años del zoológico de Independence. Al exponer su trabajo en un congreso reciente celebrado por la Acoustical Society of America relataron que su sujeto de experimentación, Lois, no percibía sonidos por encima de los 12.000 hertz pero sí captaba sonidos muy bajos, del orden de 16 hertz, a niveles de intensidad que resultaban inaudibles para el oído humano. Y llegaron a la conclusión de que la gama auditiva de los animales se halla en relación inversa con el tamaño del cráneo, que determina la distancia entre las orejas, y no con la magnitud de éstas.

La técnica de ensayo usada con Lois implicaba una recompensa cuando el elefante pulsaba correctamente con la trompa una tecla "afirmativa" o "negativa" según el sonido fuera audible o imperceptible.

McIlroy estaba allí

La mayoría de las personas sólo pisan el hospital cuando no tienen más remedio, pero los hay que se ven arrastrados por un extraño impulso que les induce a presentarse allí una y otra vez, aunque para ello hayan de simular los síntomas de enfermedad. Esta patología, por su frecuencia y extensión, se ha hecho acreedora de un nombre médico: síndrome de Munchausen (en honor de la clásica colección de cuentos que lleva por título *Las aventuras del Barón Munchausen*). Hasta la fecha, el caso histórico más representativo es el de Stewart McIlroy, cuyas andanzas recogen C. A. Pallis y A. N. Bamji en un número reciente del *British Medical Journal.*

El éxito de McIlroy en su papel de inválido imaginario se debió en gran parte a unos signos mínimos de trastornos neurológicos: una ligera atrofia muscular de la mano izquierda, una falta de reflejos en los brazos y pérdida de sensibilidad en distintas regiones de los brazos y tronco superior. Pero, por encima de todo eso, McIlroy era un genio de la simulación de síntomas. Aparecía, en varias ocasiones, padeciendo colapso pulmonar; en otras, múltiples e intensos trastornos neurológicos, fallo respiratorio agudo (que requerían una traqueotomía permanente y un tubo de plata para hablar), profundos dolores abdominales, dolores torácicos recurrentes y retención de orina. Síntomas todos ellos que se han demostrado totalmente fantaseados.

Desde 1944 hasta 1978 McIlroy pasó más de diez años en no menos de 68 centros hospitalarios distintos de las Islas Británicas: ocho hospitales de Irlanda del Norte, siete de la República de Irlanda, ocho de Escocia, dos de Gales y 45 de Inglaterra. Se ha comprobado la veracidad de 207 ingresos y hay pruebas circunstanciales de otros diez. En su dilatada carrera McIlroy costó al Servicio Nacional de la Salud la friolera de varios millones de libras esterlinas, sino de decenas de millones.

Su rica experiencia hospitalaria le permitió una notable familiaridad con la tecnología médica y le confirió una refinada habilidad en el arte de simular distintas enfermedades. Hizo de su caso un producto "artístico" de sintomatología patológica, en el que dio cabida incluso a hechos que iban más allá de su propia "dolencia"; en varias ocasiones hizo

mención de la muerte de una hermana sobrevenida por tres causas diferentes, según el lugar. Cuando un cuadro sintomatológico corría el riesgo de ser descubierto, rápidamente cambiaba a otro, y si atisbaba que su falsedad iba a quedar en evidencia inmediata, podía zafarse como el viento, curadas por ensalmo sus incapacidades parapléjica o hemipléjica. Se dio a sí mismo de alta 133 veces, agenciándoselas para huir de un posible apresamiento al emplear 22 segundos apellidos y ocho primeros apellidos en distintos centros.

Sin contar cama y pensión, McIlroy costó un riñón al contribuyente en tratamientos y observaciones. Con su habilidad usual mostraba huellas de antiguas operaciones quirúrgicas en su haber, y sus numerosas cicatrices daban fe de un sin fin de intervenciones abdominales y correcciones ortopédicas. El número de radiografías y análisis de sangre se cuentan por cientos, por miles quizá, y se sometió a no menos de 48 incisiones en columna, tres encefalogramas y múltiples mielogramas. Las notas médicas de su historia clínica ocupan 110 milímetros de espacio en el anaquel.

En 1978 McIlroy desapareció, y puesto que ha transcurrido más de un año desde que se le vio por última vez, se sospecha que esté muerto. Pallis y Bamji concluyen que el hecho de que McIlroy viviera tanto es una "prueba de la resistencia de la naturaleza humana y de la relativa seguridad de nuestros hospitales". Y añaden que su historial enseñó no poco a cuantos logró engañar.

El día en que la tierra rugió

Muchos aficionados a la vulcanología se han preguntado alguna vez qué ocurriría si se abriera un pozo en el horno de un volcán. Pues bien, se ha hecho. El pozo de sondeo número 4 (de los cinco que se construyeron) llegó hasta los 1138 metros con el fin de enlazar con un acuífero de agua caliente en el yacimiento geotérmico de Námafjall, Islandia, en 1968. Desde esa fecha hasta el 8 de septiembre de 1977, el conducto extraía el vapor producido, para lo que estaba diseñado. Pero a las 11,45 de ese día, arrojó de repente roca fundida hasta una altura de 15 a 25 metros. La columna vomitó grandes cantidades de llamas y cenizas. Esta fase de la erupción, que estuvo acompañada de un sordo bramido, duró un minuto. De los 10 a los 20 minutos siguientes, el nuevo volcán se mantuvo tranquilo. La fase final, que duró otro minuto, consistió en ráfagas de escoria (lava vesicular tosca). Luego, la chimenea volvió a su producción normal de gas.

El caso viene relatado en *Nature* por Gudrún Larsen, Karl Grönvold y Sigurdur Thorarinsson, de la Universidad de Islandia. Señalan que el yacimiento se extiende a unos nueve kilómetros al sur de la caldera (cráter volcánico) de Krafla y que el magma se introduce continuamente en depósitos magmáticos situados bajo la caldera a unos tres kilómetros de profundidad. El yacimiento de energía geotérmica se asienta en un enrejado de fallas, escenario de periódicos acontecimientos de deriva donde los segmentos de dicho enrejado de cinco a 15 kilómetros de longitud se hallan activos a causa de los movimientos de fallas, terremotos y formación de nuevos reservorios de gas. Los cinco pozos de sondeo se hallan emplazados estratégicamente para cortar las fallas a profundidades idóneas para la extracción de energía geotérmica.

Terminan su trabajo los investigadores afirmando que la erupción de la chimenea "formaba parte de un acontecimiento de deriva que ocurrió en ese día" y constituye una prueba de un movimiento horizontal a gran escala de magma basáltico desde la caldera a través de la región de fallas. Al parecer, el magma se introdujo en el pozo a través de una falla de buzamiento que cortaba la cavidad por encima del acuífero. El volumen de la erupción fue de 26 metros cúbicos y 3100 kilos de peso.

Servet y la geografía humana

Hombre de ciencia representativo del Renacimiento, enemigo de la falsa tradición, apasionado por el libre examen de los hechos, fue Miguel Servet uno de los médicos de mayor altura intelectual y preparación fisiológica de su época. Había estudiado medicina en París, en donde se doctoró. Allí trabajó con Gunther y Fernel, siendo condiscípulo de Vesalio. "Empapado en toda clase de conocimiento, superaba a casi todos por su dominio de la ciencia galénica, examinando los músculos, venas, arterias y nervios del organismo entero" (Juan Gunther). A Servet se debe, como es sabido, el descubrimiento de la circulación menor de la sangre, descrita en su obra *Christianismi Restitutio*.

Independientemente de su aportación al campo de la astronomía y astrología, en el campo estricto del conocimiento de la tierra su contribución fue notable. Pocos estudios se han publicado sobre su personalidad geográfica. El hecho de escribir sus obras en latín, lengua en la que estaba muy versado, no ha contribuido a su conocimiento. De la obra de Tolomeo, principal fuente de información geográfica durante siglos, se habían hecho numerosas versiones. La más acabada y perfecta era la del humanista alemán Pirckeymer, no exenta de errores. Sobre la base que construyera Pirckeymer, Servet realiza una labor profunda, empleándose a fondo, tanto en las bases científicas como en la metodología y forma literaria. En el prólogo hace mención a los códices griegos y latinos que le son familiares, mientras corrige erratas y añade notas y comentarios, lo que permite una mejor comprensión del texto del autor alejandrino y una lectura más clara, suave y llana.

Lejos de limitarse al camino trillado de autores anteriores, estudia con detalle los distintos países conocidos en tiempo de Tolomeo: Germania, Suavia, Tesalia, Sicilia, Cerdeña, Córcega, Bohemia, Sajonia, Moscovia, el Océano Occidental o de la Tierra Nueva, etc. También habla de Africa, del Asia, de la Tierra Santa, etc. Se refiere a ésta nombrándola como tierra seca y esteril. Ello le habría de proporcionar problemas, por cuanto Calvino, profundo conocedor de las obras de Servet, utilizó estas palabras del aragonés en sus acusaciones contra él. Cuando nombra ciudades de los más diversos países, lo hace con el nombre nuevo que se haya podido producir y en la lengua vernácula de las mismas. Al explicar los vocablos en su lengua vernácula, lo hace de tal modo que la comprensión del texto es evidente: "parece que hablamos en francés con los franceses, en alemán con los alemanes y con los españoles en español". Más tarde, añade: "Hemos visto todas estas tierras y conocemos las lenguas de todas ellas". No oculta que ha silenciado muchas ciudades que fueron descritas por Tolomeo y que en tiempo de Servet no existían porque fueron destruidas.

De todas las ciudades y regiones refiere hechos sustantivos de sus cultura y carácter, lo que proporciona a sus descripciones un notable interés. Con frecuencia utiliza el método comparativo, a imagen de Erastótenes, Polibio y Estrabón. Usa distancias, paralelos, mapas... y, por encima de todo, analiza el temperamento y los hábitos de los pueblos que estudia. En este orden, llama la atención el interés que presta al talante humano. Algunas de las comparaciones que hace resultan curiosas. Así, al hablar de los españoles, anota que su temperamento es más ardiente y seco que el de los franceses, siendo el de los franceses más frío y húmedo, y las mujeres galas más prolíferas que las españolas.

Los orígenes del primer avión motopropulsado con piloto a bordo

El avión de los hermanos Wright de 1903 no fue sólo un esfuerzo afortunado de dos mecánicos de bicicletas, sino el resultado de un programa intensivo de investigación, ingeniería y experimentación

F. E. C. Culick

En una serie de vuelos efectuados el 17 de diciembre de 1903, Wilbur y Orville Wright se transformaron en los primeros pilotos de un avión motopropulsado. Como es bien sabido, habían diseñado y construido la aeronave por sí mismos. Persiste sin embargo hasta hoy la idea de que los hermanos Wright eran esencialmente mecánicos de bicicletas, que más o menos por casualidad llegaron a su bien logrado diseño. Incluso entre los ingenieros aeronáuticos está muy extendida la ignorancia de lo que los Wright consiguieron. Siguiendo los pasos por los cuales los hermanos alcanzaron su éxito construyendo un modelo de su avión de 1903 para ensayos en túnel aerodinámico, he llegado a darme cuenta de lo importante que fueron sus logros en investigación, ingeniería y experimentación. Su trabajo merece ser presentado en un contexto histórico más rico.

En el día de sus primeros vuelos los hermanos se turnaron como piloto. Wilbur pilotó el cuarto vuelo, que fue el último y más largo de la jornada, manteniéndose en el aire durante 59 segundos. Transcurrieron casi cuatro años antes de que ninguna otra persona fuese capaz de mantenerse en el aire durante un minuto, e incluso entonces la máquina no era totalmente controlable. En esa época, los Wright habían desarrollado un avión maniobrable, capaz de volar durante más de una hora.

Ninguna otra persona entendía aún la necesidad del control lateral, ni mucho menos la función de la cola vertical, ni aún tan tarde como 1908, cuando los Wright volaron en público por primera vez. En consecuencia, ningún otro era capaz de efectuar virajes correctos. Ningún otro sabía cómo fabricar hélices correctamente y, sobre todo, ningún otro había seguido un programa comparable: acometiendo la investigación necesaria, construyendo su propio avión y volándolo por sí mismo, de forma que comprendiese el problema en toda su complejidad. Los Wright consiguieron vencer a todos los que intentaban volar porque identificaron los problemas que tenían que afrontar y resolver.

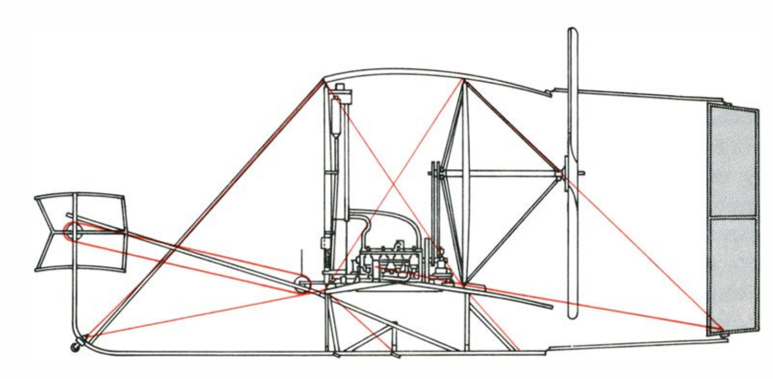
Los hermanos comenzaron su trabajo en 1899, cuando Wilbur tenía 32 años y Orville 28. Wilbur, que fue el jefe informal en gran parte del trabajo, se dio cuenta de que tendrían que seguir un programa sistemático de investigación e ingeniería. Resumió sus puntos de vista en una conferencia que pronunció en septiembre de 1901 en una reunión de la Asociación de Ingenieros del Oeste: "Los problemas del transporte terrestre y acuático se resolvieron en el siglo XIX, porque fue posible comenzar con pequeños logros y continuar gradualmente el trabajo hasta nuestro presente éxito. El problema del vuelo se dejó al siglo veinte, porque en este caso era necesario que el arte estuviese muy desarrollado antes de que fuera posible obtener cualquier vuelo de duración considerable." Cuatro años más tarde los Wright tendrían un avión práctico.

La invención del avión motopropulsado con piloto a bordo tuvo lugar parcialmente como un desarrollo natural a partir de los trabajos de otros. Las contribuciones propias de los Wright fueron, sin embargo, cruciales hasta un grado extraordinario. Y asombrosos sus logros técnicos, particularmente si se considera cuál era la situación de los esfuerzos para conseguir el vuelo humano en 1899.

Cuatro pioneros

Las personas que trataron de construir máquinas voladoras habían seguido distintas estrategias. Las observaciones del vuelo de las aves motivaron los esfuerzos, y así fue natural que las primeras ideas incorporasen siempre el uso de alas batientes. Sir George Cayley (en 1799) fue el primero en comprender que el vuelo pilotado sería más fácil de conseguir, si se separaban los medios de generar sustentación de los medios de propulsión.

Cayley está reconocido como el inventor del avión. Sus diseños y conclusiones se basaron en observaciones y experimentos cuidadosos. El concibió la configuración que hoy se considera convencional: un cuerpo o fuselaje que soporta un ala principal con colas horizontales y verticales situadas en la parte posterior. Hizo volar con éxito planeadores, uno de ellos lo suficientemente grande como



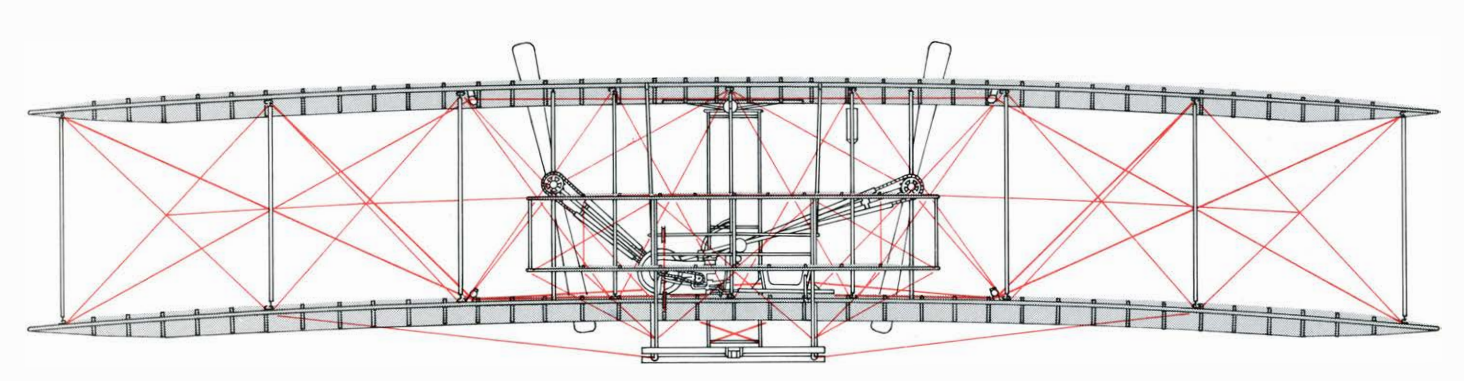
AVION DE LOS WRIGHT de 1903, que se muestra en el lado izquierdo de la figura en una vista lateral izquierda y en el lado derecho en sus vistas en planta y alzado. Fue el primer avión motopropulsado y con piloto a bordo que voló con éxito. Su rasgo más significativo era la superficie horizontal de

para llevar a un muchacho en, al menos, una ocasión. Pero Cayley nunca voló. Entre sus muchas otras contribuciones notables está la demostración de que las superficies con curvatura son mejores que las planas para proporcionar sustentación; la idea de que se puede conseguir equilibrio o estabilidad lateral dándole al ala un ángulo diedro, o sea, fabricándola con sus extremos elevados, y el concepto de utilizar una cola horizontal para lograr estabilidad longitudinal.

Lo que Cayley empezó a comprender fueron las conexiones íntimas entre la geometría de un avión, las fuerzas que actúan sobre él y su estabilidad en el vuelo. Un avión tiene un plano vertical de simetría que pasa por su eje longitudinal. Los movimientos que no desvían el avión en direcciones fuera de este plano se llaman movimientos longitudinales. El movimiento más común es el de cabeceo, en el que el morro del fuselaje se mueve hacia arriba o hacia abajo. El resbalamiento, el balance alrededor del eje longitudinal y la guiñada alrededor del eje vertical reciben el nombre colectivo de movimientos laterales.

La tendencia de cualquier objeto en movimiento estacionario de volver a su estado inicial después de una perturbación se llama estabilidad. Hay dos tipos de estabilidad, longitudinal y lateral, que corresponden a las dos clases de movimientos de un avión. En un avión moderno, la estabilidad longitudinal se debe principalmente a la superficie horizontal de cola, que normalmente está situada en la parte posterior como recomendó Cayley. La estabilidad lateral es el efecto de muchas causas, incluyendo el ángulo diedro de las alas y la cola vertical. Los diseños de las primeras máquinas voladoras incluían muchas veces una cola vertical para control direccional, análoga al timón de un barco. La verdadera función de la cola vertical como una contribución a la estabilidad no se reconoció hasta que se llevó a cabo la obra de los Wright.

El gigante de la primitiva aeronáutica



control situada por delante del ala. Las superficies verticales gemelas de cola no podían deflectarse excepto cuando el piloto alabeaba las alas para efectuar un viraje. Volando en posición tendida, el piloto conseguía el alabeo mediante cables unidos a un arnés a su cintura. El alabeo del ala, que fue un invento de los Wright, cumplía la misión que satisfacen hoy los alerones. Wilbur y Orville Wright construyeron el avión completo.

francesa, Alphonse Pénaud, fue el primero que utilizó bandas de goma retorcida para dar potencia a un modelo de avión propulsado por hélices. Independientemente de Cayley, pensó en la cola horizontal como medio para conseguir la estabilidad longitudinal. Además, comprendió el funcionamiento de la cola y dio la primera explicación de su función en el logro de un vuelo estable.

Otto Lilienthal, un ingeniero mecánico que trabajaba en Alemania, aportó contribuciones importantes que influyeron directamente sobre los Wright. En el año 1889 publicó su libro *El vuelo de las aves como base del arte de la aviación*, que contenía los resultados de muchos y largos experimentos que había llevado a cabo con ayuda de su hermano. Esta obra proporcionó los primeros datos útiles sobre la sustentación y resistencia de los perfiles aerodinámicos con curvatura. Lilienthal también experimentó con planeadores, empezando en 1891 y se convirtió en el primer piloto colgante de planeador que tuvo éxito. Ensayó diversos planeadores, tanto monoplanos como multiplanos y se mató probando un monoplano.

En los Estados Unidos el estudioso más serio de la aeronáutica durante esta época fue Octave Chanute, eminente ingeniero civil. Con sus escritos y sus propios experimentos llevó la obra de Lilienthal a los Estados Unidos. A causa de su avanzada edad (tenía 68 años en 1900) no voló él mismo. Tenía varios protegidos que ensayaron sus planeadores y sus propios diseños.

Chanute diseñó un planeador biplano que representa el principio del avión moderno. (Recientemente, el diseño sirvió de base para varios planeadores de piloto colgante.) Chanute adoptó la cola trasera de Pénaud para conseguir estabilidad longitudinal, y para mejorar la sustentación diseñó una sección de ala con curvatura similar a una ensayada por Lilienthal. La contribución más importante de Chanute es menos obvia: su elección de estructura para el biplano. Es el reticulado de Pratt que había sido patentado en 1844 como diseño para puentes de ferrocarril. Las dos alas fueron conectadas por sólidos largueros verticales que soportaban las cargas de compresión. Las cargas de tensión se transmitían mediante cables diagonales cruzados uniendo los largueros tanto en el plano lateral como en el plano anterior-posterior. El resultado fue una estructura rígida de peso ligero que se convirtió en la base de todos los biplanos con montantes exteriores.

Lilienthal y Chanute trataron de resolver el problema de la estabilidad longitudinal. Aunque fuesen capaces de diseñar sus planeadores de forma que produjeran suficiente sustentación, tuvieron gran dificultad en mantener el equilibrio en vuelo. Comprendieron que alcanzar el equilibrio significaba hacer coincidir el centro de presiones con el centro de gravedad. La dificultad aparece, sin embargo, porque el centro de presiones se desplaza, si la actitud del avión se perturba, como en el caso de una ráfaga. En un planeador de piloto colgante, el piloto puede restaurar el equilibrio desplazando su posición para mover su centro de gravedad y cambiar la actitud del avión. Si esta maniobra se realiza correcta y consecuentemente, la combinación del planeador y del piloto forma un sistema estable.

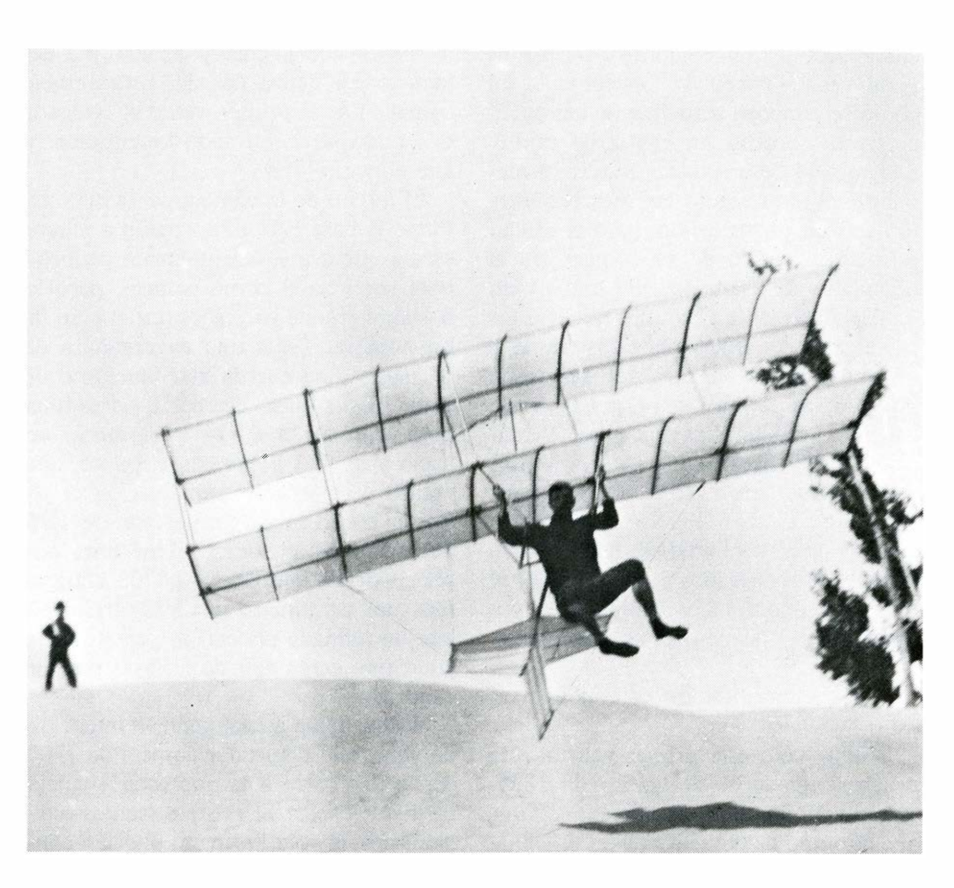
Cuando se instala correctamente una cola horizontal trasera, sirve para restaurar el equilibrio (como comprendió Pénaud). La razón es que el centro efectivo de presiones del ala y de la cola juntas permanece fijo, incluso si se perturba la actitud del avión. El avión es inherentemente estable, y el piloto no necesita



PLANEADOR PRIMITIVO, diseñado por Otto Lilienthal, a quien se ve aquí volando sobre una fábrica de ladrillos en los alrededores de Berlín en 1894. Lilienthal fue el primer piloto de lo que hoy llamaríamos un planeador de piloto suspendido. Se mató en 1896, cuando uno de sus planeadores entró en pérdida y cayó en picado. La línea irregular horizontal, que pasa cerca de las cabezas, es una grieta en el negativo de vidrio original.

desplazar su posición. Un modelo de avión es estable por esta razón. Un avión de tamaño real requiere más control, y por tanto se hace la cola movible y el piloto puede cambiar su sustentación a discreción. Con una cola móvil, un avión inestable puede ser volado por un piloto suficientemente hábil.

Los hermanos Wright empezaron a tener contacto con la aviación cuando, siendo muchachos jóvenes, les regalaron un pequeño helicóptero de juguete, parecido a uno inventado por Pénaud. Las noticias de la muerte de Lilienthal en 1896 despertaron su interés por el problema de la aviación. En mayo de 1899 Wilbur escribió al secretario de la Smithsonian Institution pidiendo bibliografía disponible sobre aviación. Entre los libros que decidió comprar figuraba el libro de Chanute *Progresos en máquinas voladoras*, escrito en 1894. Después de estudiar el libro, Wilbur escribió a Chanute, iniciando un intercambio epistolar que iba a continuar durante casi 10 años. Durante la mayor parte de este período, Chanute constituía para los Wright tanto una fuente de información como de alientos. Las cartas y diarios de los Wright representan una crónica detallada, a partir de la cual se pueden reconstruir los acontecimientos que conducen al vuelo motopropulsado.



PLANEADOR BIPLANO, diseñado en 1896 por Octave Chanute, ingeniero civil que fue un estudioso de la aeronáutica en los Estados Unidos, antes de que los Wright comenzasen su trabajo. Mantuvieron una larga correspondencia con él, tras conocer una obra suya, y les ayudó en muchos de sus primeros vuelos. El piloto de este vuelo en Indiana en 1897 fue A. M. Herring, uno de los socios de Chanute.

Comienzan los Wright

Cuando los hermanos empezaron a experimentar en serio, se habían familiarizado a fondo con la literatura aeronáutica. Comprendieron las opiniones predominantes de que la estabilidad longitudinal requería una cola de Pénaud, que se debía alcanzar la estabilidad lateral mediante un ángulo diedro de las alas y que el control direccional se podía efectuar mediante una cola vertical o timón de dirección. Sin embargo, rechazaron finalmente los tres conceptos, produciendo diseños que fueron bastante diferentes de la idea básica concebida por Cayley y aceptada en general por sus contemporáneos. Su primer avión motopropulsado, el avión de 1903, tenía una superficie horizontal delantera, las puntas de las alas caídas con diedro negativo y la cola vertical sólo se movía cuando se flexaban las alas para girar el avión. Las razones que llevaron a la configuración adoptada por los hermanos Wright cuentan la historia de su investigación e ingeniería.

Probablemente, la motivación más importante en el programa de los Wright fue que ellos mismos estaban decididos a aprender a volar. Aquella resolución influyó en todas sus decisiones. Como iban a ser pilotos, tenían que considerar cuidadosamente tanto la estabilidad como el control y la seguridad. No se sentían inclinados a construir una máquina, confiar en que funcionase y contratar a alguien para que la probase. Además, no iban a quedar satisfechos con sólo planear en vuelo rectilíneo.

La muerte de Lilienthal y un accidente similar en Inglaterra convencieron a Wilbur, al comienzo, de que un avión, para ser seguro, debía tener la superficie estabilizadora horizontal por delante del ala principal (un sistema que llegó a llamarse la configuración "canard"). Nadie en aquel tiempo comprendía la entrada en pérdida (desaparición de la sustentación), y los Wright no tenían base científica para determinar los méritos relativos de las superficies posteriores y anteriores. Mientras ellos estaban experimentando, las condiciones teóricas para la estabilidad fueron desarrolladas por otros, pero sólo después de 1903 fue ampliamente conocido que un avión se puede hacer estable longitudinalmente con una superficie horizontal anterior o posterior. La superficie anterior, que fue la característica más desacostumbrada del primer avión de los Wright, tenía para los hermanos la ventaja de que el piloto podía verla, ganando así en seguridad, al ser capaz de reaccionar con rapidez, si algo iba mal con ella en el vuelo.

En oposición a sus antecesores, Wilbur no supuso que el problema de la estabilidad lateral había sido solucionado con la invención del diedro por Cayley. "Mis observaciones del vuelo de los busardos", escribió Wilbur a Chanute, "me inducen a creer que recuperan su equilibrio lateral, cuando una ráfaga de viento los inclina lateralmente mediante una torsión de los extremos de las alas." Habiendo hecho esta observación fundamental, Wilbur tenía que solucionar el problema de inventar un método comparable de control para un avión pilotado.

Finalmente realizó la idea mediante el alabeo del ala que fue el antecesor de los alerones que cumplen la misma función en los aviones modernos. Con el alabeo, los Wright no sólo fueron capaces de estabilizar el avión en vuelo rectilíneo, sino también de hacerlo virar. Y así es como avanzaron muy por delante de sus otros competidores. Mediante el uso de una superficie horizontal móvil y del alabeo del ala, los hermanos Wright fusionaron las tareas de estabilidad y control, aportando con ello su contribución más importante al desarrollo del avión moderno.

Algunos meses después de que Wilbur escribiese a la Smithonian Institution, construyó una cometa biplana, con

cinco pies de envergadura (un pie es igual a 0,3048 metros). Las cometas, tal como se conocen actualmente (en particular las cometas en forma de cajón), habían sido desarrolladas muy recientemente. Wilbur siguió las más recientes técnicas de construcción, pero el diseño estructural básico de su cometa era el reticulado de Pratt, del que había oído hablar a través de Chanute. Eliminó los cables tensores anteriores y posteriores a fin de permitir a las alas que deslizasen una con respecto a la otra, produciendo con ello el alabeo. Las únicas partes rígidas de la estructura eran las conexiones entre la superficie de control y los soportes a que estaban ligadas. Las cuatro cuerdas de control estaban dispuestas de forma que el operador pudiese ejercer el mando longitudinal y lateral de la cometa, independiente o simultáneamente.

Experimentación en vuelo

Wilbur voló este primer vehículo de pruebas una tarde de agosto de 1899, estando Orville fuera de excursión. Aunque este tipo de cometa es algo difícil de volar, Wilbur tuvo éxito, convenciéndose de que sus ideas respecto al control longitudinal y lateral eran correctas. Descubrió también que era posible volar la cometa con la cola horizontal por delante o por detrás del ala. Esta notable cometa fue el primer vehículo volador que pudo ser controlado longitudinal y lateralmente.

El diseño de la cometa de 1899 constituyó la base para una versión a mayor escala que era suficientemente pequeña para ser volada como cometa, pero lo bastante grande para soportar el peso de un hombre. Tenía una envergadura de 17 pies y una cuerda alar (anchura) de cinco pies; su peso era de 50 libras (una libra equivale a 0,454 kilogramos) en vacío y de 190 libras con el piloto, que volaba en posición tendida. Con el piloto a bordo la carga alar era de 1,15 libras por pie cuadrado. [Una libra por pie cuadrado equivale a 0,0703 kilogramos por centímetro cuadrado.] Un velero de refinada perfección actual puede tener una carga alar de 5 libras por pie cuadrado.

Wilbur tenía inicialmente la intención de volar este planeador cometa de 1900 desde una torre a la que estaría sujeto mediante cables. Si existían vientos estacionarios, él y su hermano podrían conseguir horas de experiencia operando los mandos sin sufrir los riesgos del vuelo libre. Estimaron que, en cinco años de trabajo, Lilienthal había permanecido alrededor de 5 horas en el aire solamente, lo que apenas era tiempo suficiente para aprender a volar.

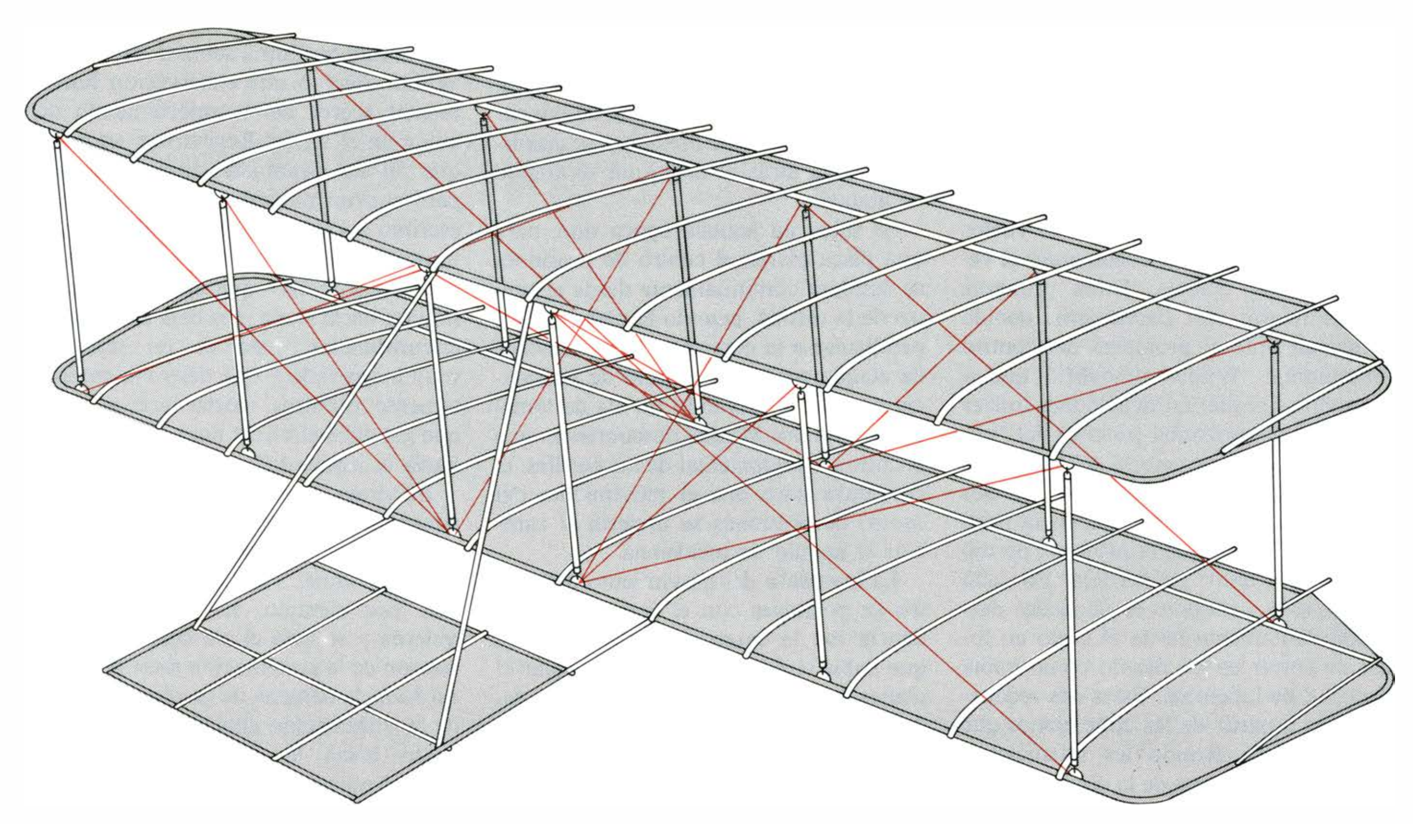
Chanute sugirió un cierto número de lugares donde los Wright podrían encontrar vientos estacionarios en la época del año en la que les sería posible experimentar. (A causa de exigencias impuestas por su negocio de bicicletas sólo podían abandonar su hogar en Dayton, Ohio, entre septiembre y enero.) Tras verificar con la oficina meteorológica las probables condiciones de viento, decidieron ir a Kitty Hawk en la costa de Carolina del Norte. Kitty Hawk era una pequeña población aislada a la que sólo podía llegarse por barco, tras un día de viaje desde Ohio. Los hermanos tuvieron que transportar su avión y todo el equipo desde la tierra firme y, además, hubieron de buscarse materiales para construir su torre.

Los Wright permanecieron menos de un mes cerca de Kitty Hawk en 1900, experimentando con su planeador entre el 5 y el 18 de octubre. La mayor parte del tiempo operaron el planeador como cometa, al principio desde una torre de 10 pies, y luego, casi siempre, desde el suelo con cuerdas sujetas a mano. El tiempo normal en vuelo libre fue alrededor de dos minutos, de los cuales aproximadamente 10 segundos fueron con hombre a bordo. Sólo Wilbur voló. Los controles trabajaban satisfactoriamente cuando el planeador volaba como cometa, pero en el tiempo limitado en que los hermanos la pilotaron no llegaron a dominar la adaptación de los controles a la operación directa.

Resulta evidente de los diarios que los primeros experimentos fueron a menudo frustrantes. Orville anotó: "Lo ensayamos con la cola por delante, por detrás y de todas las maneras. Cuando llegamos al final, Wil estaba tan confundido que no podía ni siquiera teorizar. Sólo con un esfuerzo considerable conseguí mantenerlo en el asunto de la aviación". Los vuelos eran cortos y había muchos aterrizajes violentos. Afortunadamente, la estructura era simple y fácil de reparar. Con dinamómetros unidos a los cables de retención los hermanos efectuaron mediciones de sustentación y resistencia. Determinaron la velocidad del viento con un anemómetro. Estos fueron los primeros datos cuantitativos que se obtuvieron para la construcción de una estructura sustentadora a escala no reducida.

Aunque el programa de vuelo de los Wright de 1900 fue corto, proporcionó resultados fundamentales y sirvió para "confirmar la corrección de nuestras opiniones originales". Los hermanos probaron sus técnicas de control longi-

PRIMER AVION construido por los hermanos Wright, que fue esta cometa biplano de 1899. Tenía una envergadura de 5 pies e incorporaba como estructura básica el reticulado de Pratt, que había sido diseñado para puentes de ferrocarril y adaptado al avión por Chanute. La cometa incorporaba el sistema de alabeo del ala de Wilbur. Los cuatro cables permiten un control longitudinal y lateral.



EL PLANEADOR COMETA de 1900 fue el primer avión de los Wright suficientemente grande para soportar a un hombre. Se trataba de una versión simplificada de la cometa de 1899 y tenía una envergadura de 17 pies y una cuerda (anchura del ala) de 5 pies. Los Wright lo ensayaron, principalmente como cometa, en Kitty Hawk (Carolina del Norte) en octubre de 1900. Sólo Wilbur Wright lo voló como piloto durante un tiempo total de alrededor de diez segundos. Aunque los controles del aparato funcionaban satisfactoriamente cuando los hermanos Wright volaban el avión como cometa, Wilbur descubrió, en su breve intento de pilotarlo, que no podía dominar la adaptación de los controles del planeador a la operación directa.

tudinal y lateral; se confirmaron aún más en su creencia de que tenían que practicar el vuelo y, a partir de sus mediciones, aprendieron que volando en una posición de piloto tendido se reducía la resistencia de la máquina.

Resultados de la experimentación

No todas sus conclusiones fueron tan positivas. En esta ocasión Wilbur decidió no fabricar las alas con ángulo diedro. Aunque la configuración contribuía a la estabilidad con aire en calma, los extremos ligeramente levantados de las alas tendían a acentuar el efecto de las ráfagas laterales. Como los Wright disponían de control lateral a través del alabeo del ala, no necesitaban el efecto estabilizante del diedro.

Con mucho, los resultados más decepcionantes de los experimentos de 1900 fueron los valores de sustentación y resistencia que obtuvieron los hermanos. La resistencia de la estructura, cuando no se generaba sustentación, era mucho menor de la que habían esperado. Sin embargo, para un ángulo dado entre el ala y el viento, la resistencia total era mayor y la sustentación menor de la que habían anticipado. (Los Wright trabajaban a partir de datos sobre una placa plana perpendicular a la corriente, obtenidos por John Smeaton en 1759 y de mediciones para perfiles con curvatura efectuadas por Lilienthal.) Los hermanos observaron en sus ensayos en vuelo que el planeador operaba a un ángulo de incidencia mucho mayor del que habían esperado. La sustentación aumenta con el ángulo de incidencia y debe ser igual al peso del planeador en vuelo estacionario. Para un ángulo de ataque dado, las alas de los Wright producían menos sustentación de la que habían previsto los hermanos basándose en los datos de Lilienthal.

Este fue un resultado sorprendente, pero tenían confianza en sus propios datos y consideraron tres posibles explicaciones: (1) el tejido del revestimiento no era suficientemente impermeable al viento, (2) la curvatura de su perfil (1:22) era demasiado pequeña para que fuesen aplicables los datos de Lilienthal o (3) los datos de Lilienthal eran incorrectos. Se inclinaron por la segunda razón y diseñaron su planeador siguiente, que iba a ser volado en otoño de 1901, con un perfil de mayor curvatura.

Con la interpretación de sus resultados de 1900, los Wright dejaron de depender del trabajo de otros. A partir de entonces, su progreso fue enteramente resultado de sus propios descubrimientos. Empezaron a trabajar en su planeador de 1901 durante el invierno. Para obtener más sustentación aumentaron la envergadura del ala a 22 pies. Con una cuerda de 7 pies la superficie alar total era de 290 pies cuadrados; la superficie del estabilizador era de 18 pies cuadrados. Este era, con mucho, el mayor planeador que nadie hubiese tratado nunca de hacer volar. El peso de la estructura se había transformado en 98 libras y la carga alar con un piloto era de 0,78 libras por pie cuadrado. Los únicos cambios adicionales de importancia consistían en el incremento de la curvatura de 1:22 a 1:12 y en modificaciones de los controles de forma que el piloto pudiese operar con mayor facilidad el estabilizador y el mando de alabeo del ala. Aunque el planeador no tenía un aspecto muy diferente al del modelo de 1900, era en algunos aspectos el instrumento más importante de investigación que tuvieron los Wright. En efecto, el año 1901 representó un momento crucial para casi todos sus logros posteriores.

Ampliación de los ensayos

El programa de experimentación se estaba haciendo cada vez mayor. Los hermanos construyeron un cobertizo de madera en Kill Devil Hills, a unos 6,5 kilómetros de Kitty Hawk, donde las grandes dunas de arena proporcionaban buenas plataformas de lanzamiento. Debido a presiones de su negocio tuvieron

que efectuar la experimentación en vuelo desde el 10 de julio hasta el 20 de agosto, antes de lo que habían planificado inicialmente.

Según sus diarios, Wilbur hizo todos los vuelos de 1901. En el primer día de ensayos, el 17 de julio, efectuó 17 vuelos de planeo sin haber probado antes el vehículo como cometa. Estos primeros vuelos fueron casi desastrosos, debido principalmente al problema del control longitudinal. Wilbur descubrió que a menudo se requerían deflexiones totales del control horizontal anterior (el "canard") y que la trayectoria del planeador presentaba severas oscilaciones. En dos ocasiones, el ángulo de incidencia llegó a ser tan elevado que la máquina perdió velocidad y entró en pérdida. En cada una de estas ocasiones el planeador descendió lentamente hacia el suelo en lugar de entrar en un picado como había hecho el de Lilienthal. Estas dos recuperaciones a partir de las condiciones que Wilbur había temido les satisficieron con respecto al valor de la configuración "canard". Desde entonces y hasta 1910 no tuvieron interés en ensayar la configuración de la cola posterior.

Wilbur se encontró con el problema de determinar por qué el planeador se comportaba de una forma tan errática. Una reducción del área de superficie de control de 18 a 10 pies cuadrados no aportó ninguna mejora sensible. Los hermanos volaron entonces el planeador como cometa, y en una tarde descubrieron que el centro de presiones se desplazaba hacia atrás al disminuir el ángulo de ataque.

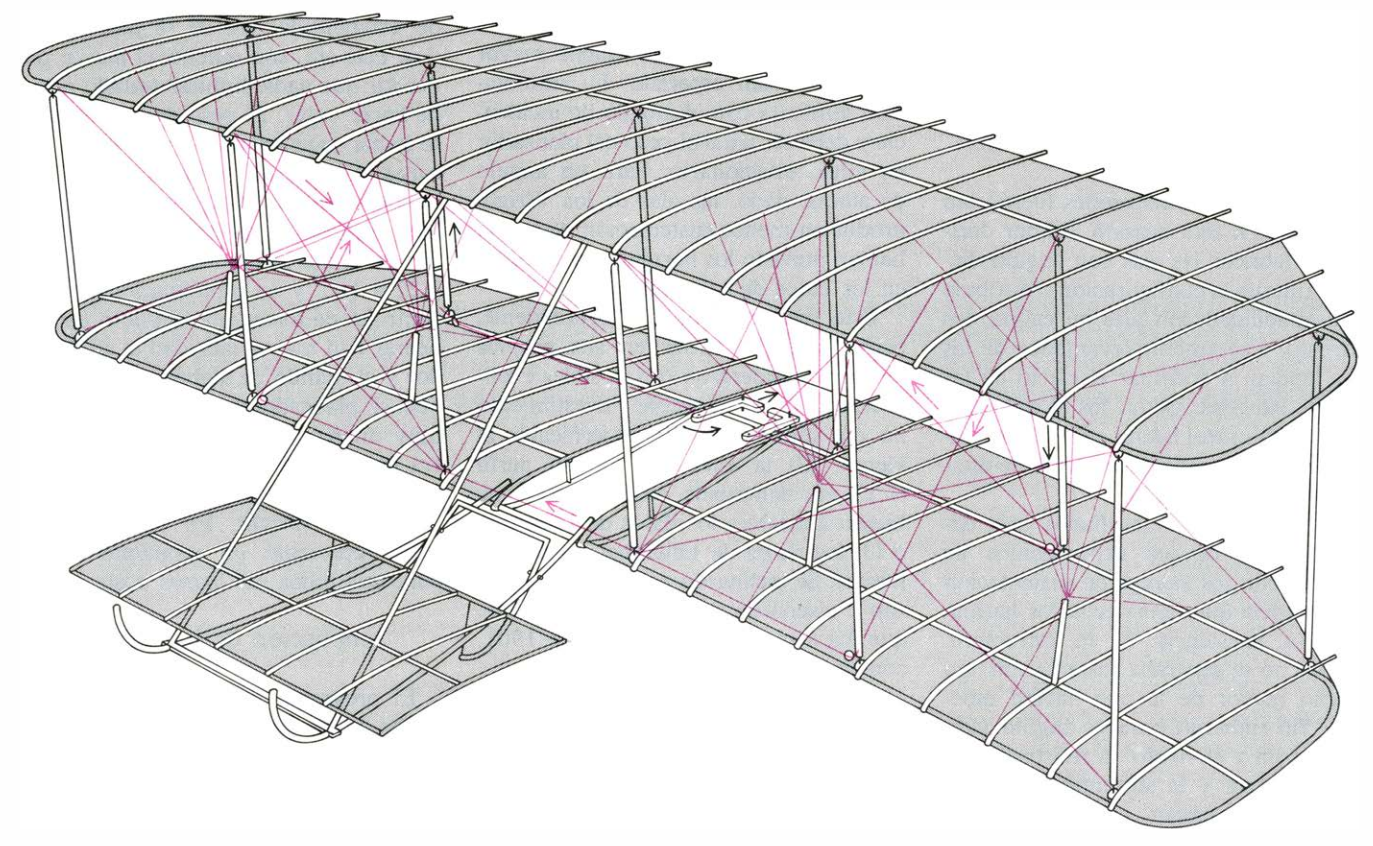
Se sabía en aquella época que, para una placa plana, el centro de presiones se desplaza continuamente desde el centro de la misma, cuando la placa es perpendicular a la corriente, hasta el borde de ataque cuando el ángulo de incidencia es nulo. Ensayando un ala de tamaño natural, los Wright descubrieron una propiedad fundamental de los perfiles: la curvatura hace que el movimiento del centro de presiones se invierta al cambiar el ángulo de incidencia.

La inversión del movimiento del centro de presiones con el ángulo de incidencia fue la causa de las dificultades que había tenido Wilbur en controlar el planeador. La solución consistió en reducir la curvatura de las alas. Los hermanos volvieron a la curvatura del planeador de 1900, instalando largueros y fijando alambres tensores para desplazar hacia abajo el punto medio de las costillas de las alas. Los resultados fueron inmediatos y satisfactorios. Wilbur hizo alrededor de 30 planeos tras la modificación; el más largo de los mismos duró 17,5 segundos y la mayor distancia recorrida fue de 390 pies.

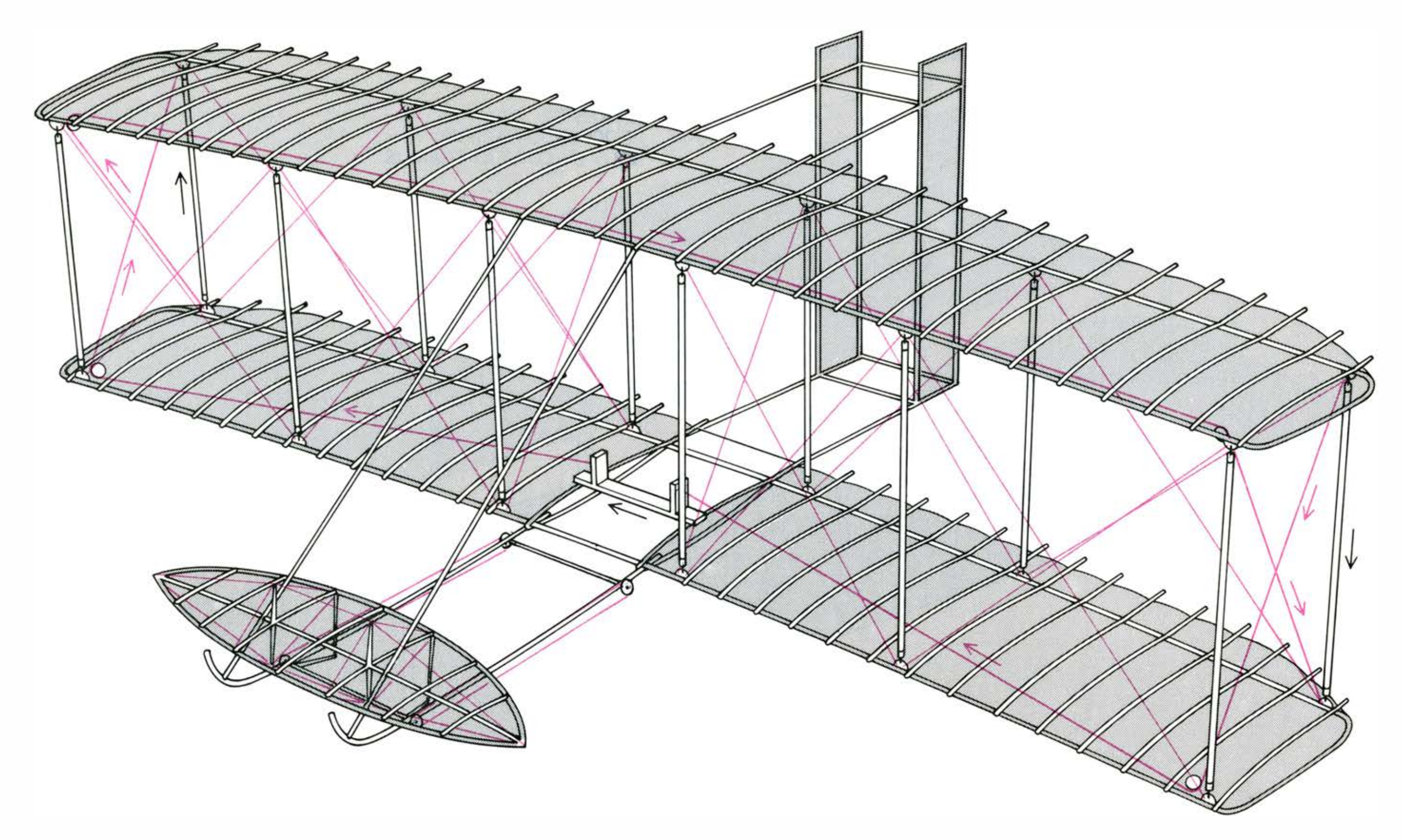
Durante la última semana de ensayos Wilbur efectuó una observación fundamental acerca del comportamiento del avión en el viraje. Registró en su diario que "el ala levantada parece retrasarse, pero al principio sube". El 22 de agosto escribió a Chanute: "La pasada semana transcurrió sin grandes resultados aunque comprobamos que nuestra máquina no vira hacia el ala más baja en todas las circunstancias. Esto es un resultado muy inesperado y que destruye completamente nuestras teorías acerca de lo que produce el viraje hacia la derecha o hacia la izquierda".

El viraje de un avión no se efectúa controlando con el timón de dirección, sino mediante un balance alrededor del eje longitudinal. Para virar hacia la derecha, por ejemplo, se eleva el ala izquierda y se baja el ala derecha. La dirección de la sustentación total se inclina así hacia la derecha de la vertical y parte de la sustentación actúa para acelerar el avión hacia la derecha. Fueron los Wright quienes descubrieron el método correcto de viraje, un descubrimiento hecho posible por su invención del alabeo del ala para control lateral.

Cuando Wilbur estaba practicando el viraje descubrió lo que hoy día se conoce por el nombre de guiñada adversa. Como inicialmente la sustentación del ala que se eleva es mayor que la del ala



PLANEADOR DE 1901. Fue, con mucho, el de mayor tamaño que nadie hubiese intentado volar nunca. Tenía una envergadura de 22 pies y una cuerda de 7 pies; el peso sin piloto era de 98 libras. En los ensayos en Kitty Hawk, con Wilbur como piloto, el planeador siguió al principio una trayectoria que presentaba severas oscilaciones. Los Wright modificaron el avión para reducir la curvatura que habían añadido a las alas desde el año 1900.



PLANEADOR DE 1902. Este aparato les proporcionó a los Wright la mayor parte de su experiencia de vuelo. Fue el primero de sus aviones que tuvo la doble cola vertical, que se añadió para compensar la tendencia del avión a resbalar en el viraje, y que incluyó el arnés de cintura. El avión tenía una envergadura de 32 pies y una cuerda de 5 pies. Los Wright habían aprendido de sus ensayos en túnel aerodinámico que un ala larga y estrecha es más efectiva que una más corta de la misma superficie. Otra característica de este planeador es que las alas están montadas con diedro negativo.

que desciende, lo mismo ocurre con la resistencia. La diferencia en fuerzas de resistencia tiende a hacer virar el avión en la dirección opuesta a la deseada cuando se inicia el movimiento de balance. Las observaciones de Wilbur condujeron a la instalación de la cola vertical en el planeador de 1902.

Como Wilbur era, a un tiempo, teórico, inventor, constructor y piloto de pruebas, fue capaz de hacer en tres semanas de vuelo dos descubrimientos cruciales: uno, la influencia de la curvatura sobre el movimiento del centro de presiones y, el otro, la guiñada adversa. Sus respuestas completaron esencialmente la configuración del avión.

El túnel aerodinámico

Los experimentos de 1901 convencieron a los Wright de que los datos de Lilienthal para perfiles con curvatura eran incorrectos. Mediciones más cuidadosas habían confirmado los resultados, obtenidos en 1900, de que la sustentación predicha era demasiado grande. Los hermanos desarrollaron una larga serie de ensayos en un túnel aerodinámico, comenzando en otoño de 1901.

No fue el suyo el primer túnel aerodinámico, ni siquiera el primero de los Estados Unidos. Probablemente tuvieron la idea a partir de artículos en el *Aeronautical Annual*, que fue publicado por James Means en Boston en los años 1895-97. La contribución de los Wright fue dirigida, como era típico en ellos, directamente a los aspectos de interés. Registraron datos muy exactos para muchos perfiles y formas de ala. No efectuaron medidas absolutas, sino que prefirieron comparar las fuerzas que actuaban sobre un perfil con las que actuaban sobre una placa plana orientada perpendicularmente a la corriente. Como resultado de estos ensayos, hacia mediados del invierno tenían todos los datos para diseñar su avión de la próxima década.

Los Wright decidieron, también basándose en sus propios datos, que las medidas de Lilienthal no eran tan inexactas como inicialmente habían supuesto. Para las predicciones de sustentación de sus planeadores habían utilizado una constante (relacionada con la resistencia de la placa plana), que dedujeron a partir de datos publicados por Smeaton. Comparando los resultados de túnel con las medidas de sustentación en sus planeadores, los Wright determinaron que el valor de la constante era erróneo. Este error les había llevado a sobreestimar la sustentación en un 40 por ciento.

El 28 de agosto de 1902 los Wright llegaron nuevamente a Kill Devil Hills con otro planeador. Este modelo tenía una envergadura de 32 pies y una cuerda de 5 pies, de forma que su superficie total y su carga alar eran aproximadamente las mismas que las del planeador de 1901. Los hermanos aprendieron de los ensayos en túnel aerodinámico que un ala larga y estrecha, con un alargamiento elevado, es más eficiente que otra con la misma superficie, pero de envergadura más reducida. Su elección de la forma de perfil se basaba también en los ensayos en túnel aerodinámico.

La máquina de 1902 fue la primera que incluía el arnés de caderas que diseñaron los hermanos para facilitar al piloto en posición prona el poder manejar los cables de alabeo, desplazando sus caderas lateralmente. La deflexión de la superficie horizontal se controlaba mediante una palanca que se manejaba con la mano izquierda.

El cambio más evidente e importante de configuración consistió en añadir una doble cola vertical. Tenía una superficie total de 12 pies cuadrados. Esta superficie estaba montada rígidamente en el planeador de 1902 y nunca se pensó en utilizarla para el control direccional. Los Wright habían calculado correctamente que la cola vertical era precisa para compensar la guiñada adversa.

SISTEMA DE ALABEO del ala inventado por los hermanos Wright como medio para realizar el viraje. En la parte superior se muestra una vista frontal de las alas. Cada uno de los ocho tramos entre los montantes está sujeto mediante un reticulado de cables, mientras que sólo cuatro de los tramos en la parte posterior poseen dicho reticulado. De esta forma los bordes exteriores de las alas traseras podían ser desplazados (*abajo*) mediante cables unidos al arnés de cintura. El mismo sistema de cables actuaba también las bielas que accionaban la cola vertical. Para efectuar un viraje, por ejemplo a la derecha, el piloto alabeaba las alas de la derecha hacia arriba y las alas de la izquierda hacia abajo. (Dibujos de Dan Todd.)

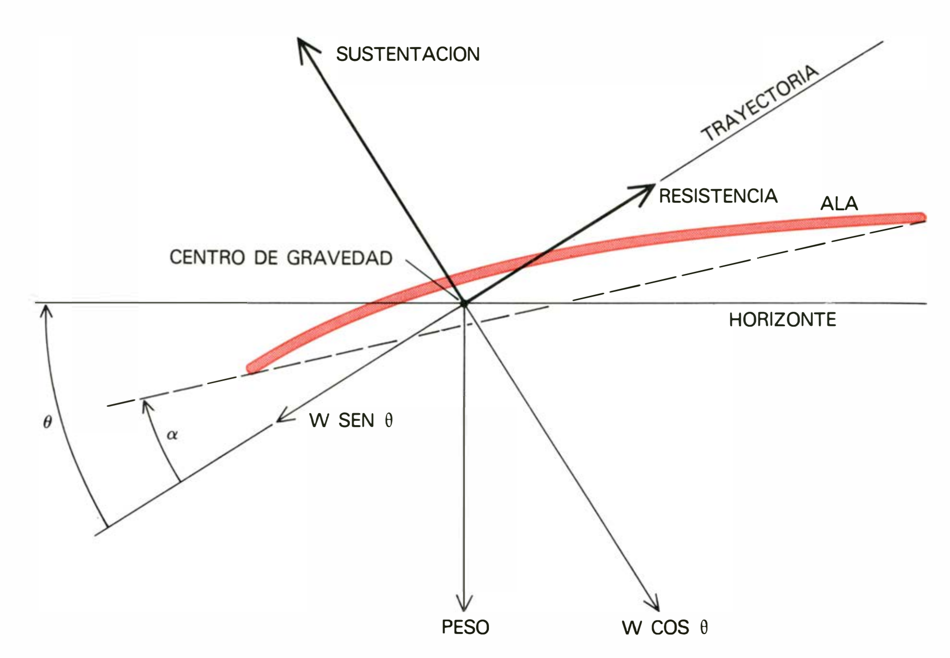
Los ensayos del planeador como cometa demostraron a los hermanos que tenían en él un avión mejor. Descubrieron que, con la nueva ala, la relación sustentación-resistencia había mejorado mucho y que las dificultades con el control longitudinal eran menores. Además, la cola vertical fija actuaba a su vez para reducir la tendencia a resbalar en el viraje.

Las pruebas iniciales de planeo mostraron una reacción excesiva a ráfagas laterales y por eso se montaron las alas con diedro negativo. Los ensayos siguientes demostraron que con el planeador de 1902 el piloto tenía incluso más dificultades con el control lateral que con la máquina de 1901. Cuando el planeador quedaba expuesto a una ráfaga lateral, tenía una tendencia incontrolable a oscilar en guiñada.

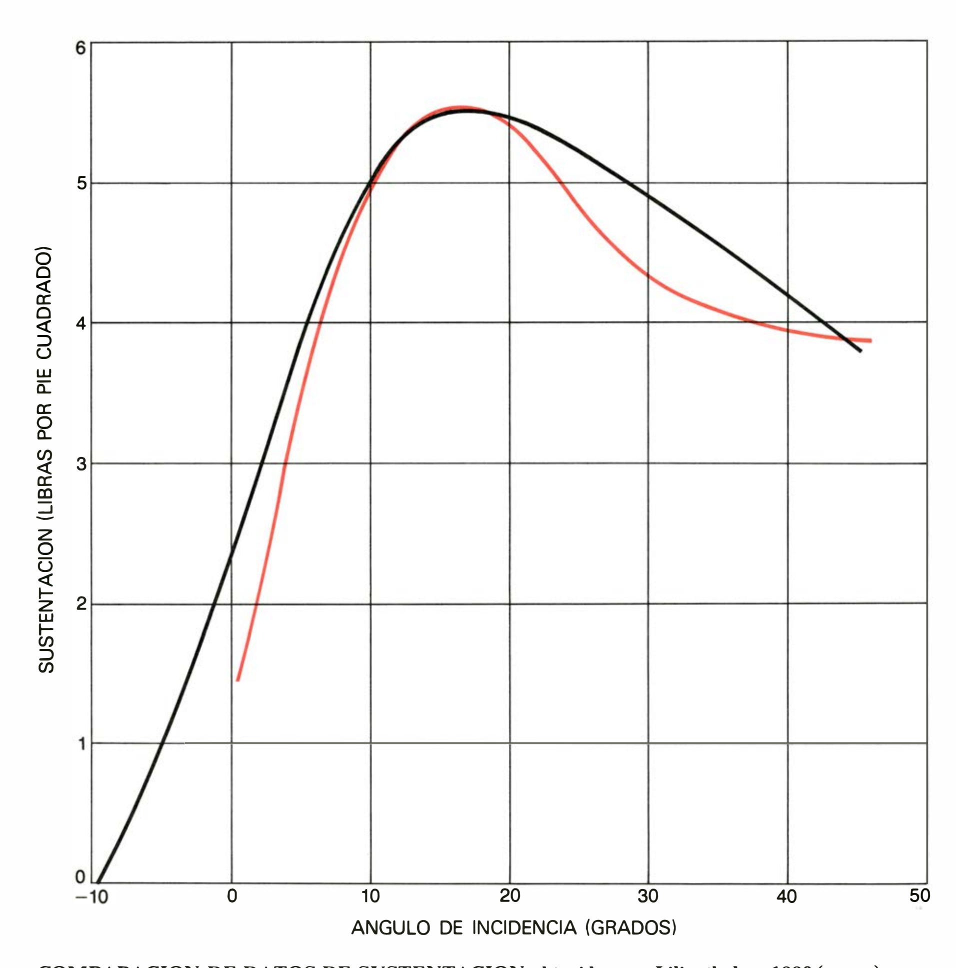
Problemas mucho más serios surgieron en varios de los virajes. Ambos hermanos experimentaron el principio de una barrena en la dirección del giro. La secuencia de entrada en pérdida-barrena en el viraje ha sido estimada como causa de la pérdida de control en un 70 por ciento de los accidentes de aviación. Los hermanos Wright fueron los primeros que la descubrieron. Empezaban un giro, digamos hacia la derecha, alabeando los bordes de salida, hacia abajo en el ala izquierda y hacia arriba en el ala derecha. La fuerza que actuaba sobre la cola vertical tendía también a virar el avión hacia la derecha, actuando así en contra de la guiñada adversa. Si el giro se hacía demasiado inclinado, había que invertir el alabeo. Esta maniobra daba lugar a guiñada adversa actuando en la misma dirección que el efecto del giro sobre el estabilizador vertical. Por eso, durante un tiempo corto, el avión viraba aún más hacia la derecha; esto supone el comienzo de la barrena. También podía ocurrir que el excesivo alabeo hacia abajo de las alas en la parte interior del viraje diese lugar a que entrasen en pérdida y cayesen. Varias veces los hermanos se encontraban con que esta serie de acontecimientos terminaba cuando el ala que estaba en la parte interior del viraje caía en la arena y el avión pivotaba y se detenía.

Fue idea de Orville corregir el comportamiento sustituyendo la cola doble fija por una móvil simple. Wilbur sugirió que su operación se ligase a la de los cables de alabeo; en estas condiciones no se podía actuar por separado. No obstante, la inversión del alabeo también invertía la tendencia de guiñada producida por la tendencia a entrar en barrena.

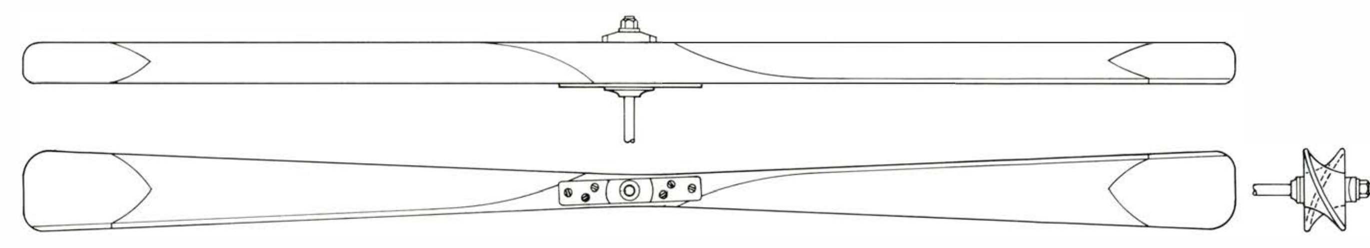
Estos varios ajustes completaron la configuración y el sistema de control



FUERZAS QUE ACTUAN SOBRE EL ALA de un planeador. Comprenden el peso (*W*), la sustentación y la resistencia. En el procedimiento empleado alrededor de 1900 para diseñar la superficie del ala, se suponía ésta en vuelo estacionario (aquí hacia la izquierda) a una velocidad (*V*) según una trayectoria que formaba el ángulo θ por debajo del horizonte. Una pequeña parte del peso (*W sen* θ) actúa para tirar del planeador en adelante, equilibrando exactamente a la resistencia. La sustentación se definió como perpendicular a la trayectoria, soportando la otra componente (*W cos* θ) del peso. La sustentación y la resistencia son funciones del ángulo de incidencia (α), que aparece a la izquierda. (En color, el ala.)



COMPARACION DE DATOS DE SUSTENTACION obtenidos por Lilienthal en 1889 (*negro*) y por los Wright en 1901 (*color*). Los Wright tuvieron que enfrentarse con el hecho de que sus primeros ensayos en vuelo intensivos efectuados con el planeador de 1900 mostraban que sus alas generaban menos sustentación de la que los datos de Lilienthal les habían llevado a esperar. La primera solución para corregir la situación consistió en aumentar la curvatura de las alas; después efectuaron ensayos en túnel. Al final los hermanos descubrieron que sus resultados eran muy próximos a los dados por Lilienthal con un ángulo de incidencia del ala entre 5 y 10 grados y que el error tenía otro origen.



DISEÑO DE HELICE DE LOS WRIGHT, que desarrollaron por sí mismos tras descubrir que los datos publicados sobre hélices marinas ofrecían poca información útil. Estaba basado en el principio de que un pequeño segmento de la pala podía tratarse como una sección de ala. Los datos marinos sugerían que el rendimiento más elevado para una potencia dada se obtendría haciendo pasar la mayor cantidad posible de aire a través del disco de la hélice. Este concepto explica el diámetro más bien alto (8,5 pies, unos 2,59 metros) de las dos hélices del avión de los Wright de 1903.

que los hermanos emplearon en el avión de 1903. Sólo a finales de 1904 desconectaron la cola vertical y lograron por fin un control independiente de cabeceo, balance y guiñada. Prácticamente todos los aviones modernos dependen del mismo sistema de control.

Proporcionando potencia

Las perspectivas para construir una máquina voladora parecían mucho más favorables después de la invención del motor de combustión interna a finales del siglo XIX. Antes que los Wright, distintas personas se habían ocupado del problema, y su contemporáneo, Samuel P. Langley, secretario del Smithsonian, llegó a punto de ensayar un avión pilotado motopropulsado a finales de 1903. En dos pruebas, que tuvieron amplia publicidad, en las que el avión fue lanzado mediante catapulta desde una embarcación en el río Potomac, no consiguió volar y cayó en el río. El progreso de los Wright no se conocía públicamente, pero ellos eran conscientes de que existía una competencia.

Orville y Wilbur habían empezado a trabajar en su avión en otoño de 1902. Su plan era el de instalar motor y hélices en una versión mejorada del planeador de 1902. Suponían que podrían adaptar un motor de automóvil y que la literatura técnica naval les proporcionaría la base para el diseño de hélices eficientes.

Después de haber intentado conseguir sin éxito de la industria un motor adecuado, decidieron diseñarlo y construirlo por sí mismos. Esto no fue una decisión especialmente difícil para ellos porque ya habían construido un motor de gasolina para accionar maquinaria en su taller de bicicletas. En marzo de 1903 obtuvieron lo que deseaban: un motor adecuado que desarrollaba 12 caballos de potencia y pesaba 15 libras por caballo. No habían tratado de crear un motor excepcional, y el resultado fue muy inferior al motor construido para Langley por su asistente de ingeniería, Charles Manly.

Los hermanos tuvieron entonces la sorpresa de descubrir que la teoría de hélices disponible sólo proporcionaba un resultado útil: para una potencia dada, el mejor rendimiento se consigue haciendo que el flujo de aire a través de la hélice sea el mayor posible. A esto se debe el diámetro bastante elevado de sus dos hélices: 8,5 pies. La consecuencia más importante de este descubrimiento fue que estaban obligados a elaborar sus propios métodos para diseñar hélices.

No existía teoría alguna para el diseño detallado de ningún tipo de hélice. Los Wright emplearon sus conocimientos sobre perfiles aerodinámicos para desarrollar lo que más tarde se conocería como teoría del elemento de pala, basada en la idea de tratar un pequeño segmento de la pala como sección de un ala. Como la hélice está en rotación y el avión se desplaza, cada elemento de esta ala está expuesto a una velocidad distinta y por ello genera fuerzas de sustentación y resistencia diferentes. La tracción de la hélice y la potencia absorbida por la misma se obtienen sumando las contribuciones de sustentación y resistencia de todos los elementos de pala.

Los análisis de los Wright les enseñaron a diseñar hélices que tenían un rendimiento del 70 por ciento o aún mejor en una época en que las hélices de otros investigadores no conseguían más del 50 por ciento. Con estas hélices los hermanos fueron capaces de propulsar su avión de 1903 pese a su motor relativamente modesto. El trabajo sobre diseño de hélices fue su mayor logro analítico; mas, para ellos, lo único importante estaba en las consecuencias prácticas. Nunca publicaron sus ideas y análisis, y la teoría moderna de hélices se basa en otras fuentes.

El primer vuelo

El motor y las hélices se instalaron en un avión que tenía la misma configuración que el planeador de 1902, pero que era mayor y estaba mejor construido. Cualquiera que examine el avión o los planos del mismo, tiene que quedar impresionado por la maestría de los Wright en las técnicas basadas en la madera, el alambre y el entelado. El avión tenía una envergadura de 40 pies, una cuerda de 6,5 pies y una superficie total de 510 pies cuadrados. El peso con piloto era de 750 libras, lo que daba lugar a una carga alar de 1,5 libras por pie cuadrado.

Wilbur y Orville llegaron a Kill Devil Hills el 25 de septiembre de 1903, esta vez para su estancia más larga, que fue de tres meses. Reconstruyeron su cobertizo de 1902 y construyeron uno nuevo (44 por 16 pies) para el avión. Cuando el tiempo era propicio, perfeccionaban sus habilidades como pilotos con el planeador de 1902. Pasaron el resto del tiempo montando el nuevo avión y haciendo ensayos previos con él.

En octubre los hermanos se enteraron de la noticia del fracaso del primer ensayo de Langley. La noticia pudo haber sido alentadora, pero sus propias posibilidades de éxito eran escasas y ellos lo sabían. Su avión pesaba 25 libras más de lo esperado. La máquina no había sido montada antes del viaje a Kitty Hawk, de forma que el motor y las hélices no se habían aún operado en la misma. En el primer ensayo el motor no funcionó satisfactoriamente y después fallaron dos soportes del eje de la hélice. Los enviaron a Dayton para su reparación el 5 de noviembre.

Hasta que les devolvieron los ejes el 20 de novimbre, los hermanos podían hacer poco. El tiempo transcurría rápidamente y empezaba a resultar escaso, sobre todo porque los Wright sabían que Langley estaba preparando otro ensayo. Las condiciones meteorológicas se estaban haciendo difíciles, ya que de vez en cuando la temperatura descendió por la noche por debajo de cero grados, pero el 23 de noviembre se habían hecho ya bastante ensayos como para medir el empuje total: 132 libras. Como los hermanos habían estimado que la resistencia sería de 95 libras, confiaron ahora en que su máquina despegaría y volaría.

El 28 de noviembre, con el motor en marcha para probar los instrumentos, se rompió un eje de la hélice. Tenía que ser

sustituido, y por ello Orville volvió a Ohio para hacer fabricar dos nuevos de material más fuerte. En el tren de vuelta de Dayton a Kitty Hawk leyó informes en el periódico indicando que la máquina voladora de Langley había quedado destruida en el lanzamiento el día anterior, el 8 de diciembre. La competición había terminado.

Una vez terminadas las reparaciones, se preparó el avión para ensayarlo el 14 de diciembre. Wilbur ganó, al echar a suertes con una moneda, y fue piloto en el primer vuelo. El carril de lanzamiento estaba ligeramente inclinado, de modo que la gravedad ayudase al lanzamiento a efectuar contra una leve brisa. El avión despegó con viento lateral suave, pero pronto entró en pérdida, cayendo al suelo a 60 pies del extremo del carril. Las superficies de control anteriores y su estructura de soporte resultaron dañadas, pero pudieron ser reparadas fácilmente.

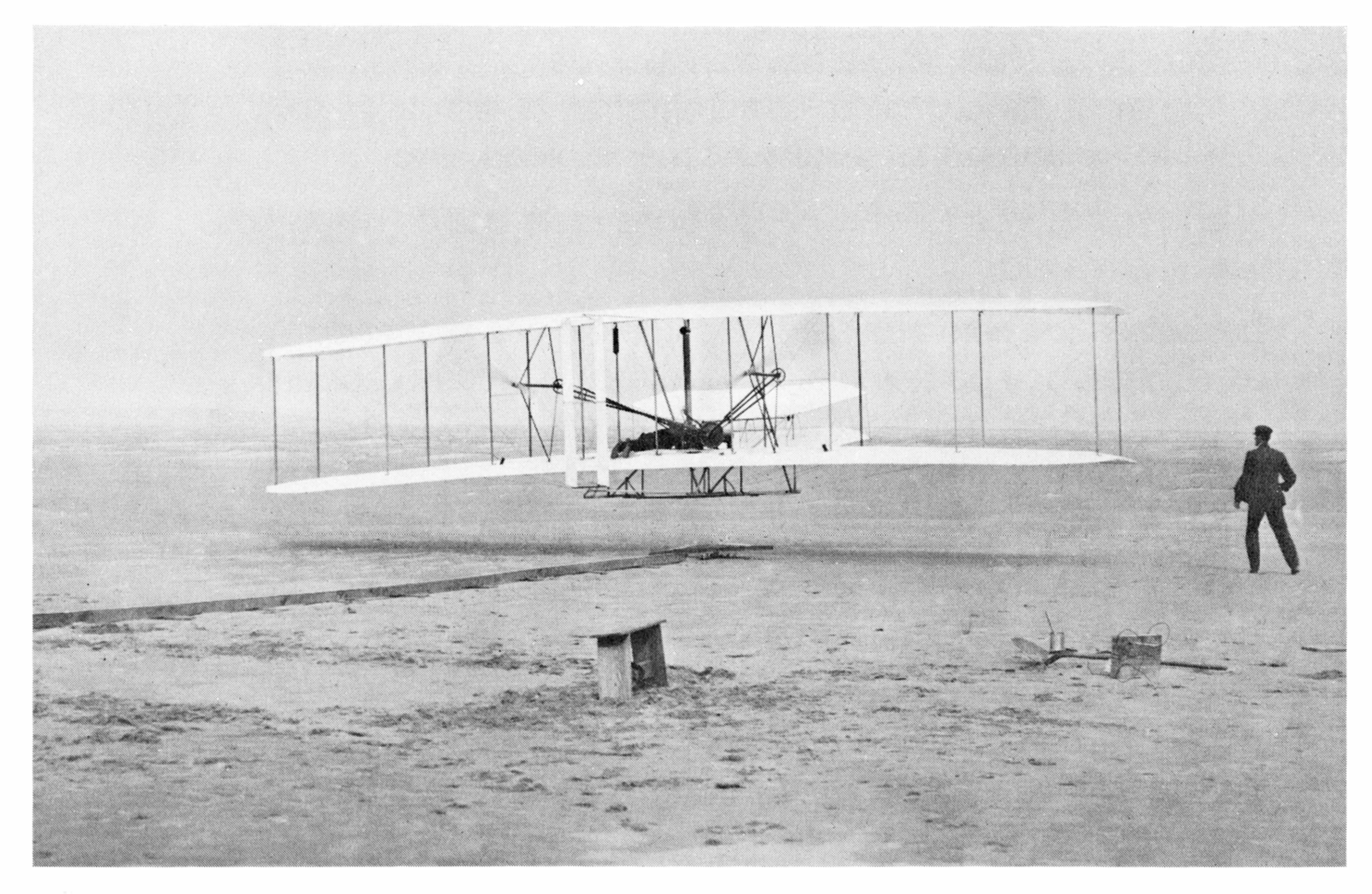
Orville pilotó el avión en el segundo ensayo que tuvo lugar el 17 de diciembre. El viento era estacionario, con una velocidad de 27 millas por hora, lo que representaba una fracción peligrosamente alta de la velocidad de vuelo del avión de 30 a 35 m.p.h. Wilbur guió el ala de estribor durante el lanzamiento, con el carril colocado ahora horizontal sobre la arena, y Orville despegó. En su vuelo recorrió una distancia de 120 pies a una velocidad media de 7 m.p.h. Este bien logrado vuelo es motivo de una de las fotografías más emocionantes y conocidas de la aviación, que muestra el avión volando a pocos pies sobre el suelo y a Wilbur corriendo a su lado.

Cada hermano efectuó dos vuelos aquel día, antes de que el avión fuese arrastrado sobre la arena por el viento y resultase demasiado dañado como para poder ser reparado con facilidad. Después de permanecer un total de 98 segundos en el aire, el primer avión motopropulsado y pilotado no volvió a volar nunca. Se exhibe ahora en el Museo del Aire y del Espacio de la Smithsonian Institution.

Durante dos años más, los hermanos Wright trabajaron solos para perfeccionar el nuevo avión. Su avance no fue conocido por la opinión pública. En la comunidad aeronáutica, Chanute era el único no escéptico respecto a su pretensión de poseer la primera máquina voladora que constituía un verdadero éxito. En 1906 les fue concedida a los Wright una patente de los Estados Unidos para su sistema de control lateral, incluyendo la deflexión de las superficies de las alas y el movimiento de la cola vertical.

Sólo después de haber completado acuerdos de concesión de licencias satisfactorios en los Estados Unidos y en Europa, estuvieron los hermanos dispuestos a exhibir su avión (versión mejorada del avión de 1903) en público. Wilbur hizo el primer vuelo público cerca de Le Mans en Francia en agosto de 1908; pocos días después, Orville efectuó una demostración en Fort Myer, Virginia. Las habilidades aeronáuticas de los hermanos sorprendieron al mundo y los tenaces Wright se hicieron internacionalmente famosos.

Sus aviones de 1908 tenían la misma configuración básica que el avión de 1903, pero el diseño no pudo sobrevivir. En menos de dos años los aviones con colas posteriores mostraron rendimientos superiores, y los Wright se vieron obligados a abandonar el "canard". Después contribuyeron poco al desarrollo del avión, pero sus logros en el período de 1899 a 1908 siguen siendo únicos en la historia.



PRIMER VUELO del avión de 1903. Fue fotografiado en Kill Devil Hills cerca de Kitty Hawk por John T. Daniels, de la Estación de Salvamento de Kill Devil, utilizando la cámara de los Wright. La fecha fue el 17 de diciembre de 1903. Orville es el piloto y Wilbur está corriendo a su lado. En este vuelo, uno de los cuatro realizados por los Wright en ese día, se recorrió una distancia de 120 pies a una velocidad media de 7 m.p.h. con respecto al suelo. Wilbur efectuó el vuelo más largo y último del día, permaneciendo en el aire durante 59 segundos. Tras un tiempo total de 98 segundos en ese día, el avión nunca volvió a volar, siendo sustituido por modelos mejorados. El avión se expone en el Museo del Arte y del Espacio de la Smithsonian.

Compartimentos en el desarrollo de los animales

La mosca del vinagre, y quizá también otros animales, está construida a partir de unidades llamadas compartimentos, grupos de células cuyo programa específico de desarrollo está dirigido por ciertos genes

Antonio García-Bellido, Peter A. Lawrence y Ginés Morata

El desarrollo de un animal es un milagro tan cotidiano que lo consideramos como trivial y no nos sorprendemos de su exactitud. Y, sin embargo, es un proceso mucho más complejo que cualquiera de los diseñados por el hombre, los cuales, contienen frecuentemente además errores de construcción y ensamblaje cometidos durante el proceso que va desde el diseño original hasta el producto final. En el caso de los seres vivos, el diseño original está contenido en el ADN de los cromosomas en forma de secuencias de nucleótidos, que son las subunidades moleculares del ADN. Ahora bien, ¿cómo se traduce esta información genética unidimensional del ADN en estructuras tridimensionales como son los diversos órganos de los seres vivos? La búsqueda de los principios que controlan estos procesos de morfogénesis es uno de los objetivos fundamentales de la biología del desarrollo.

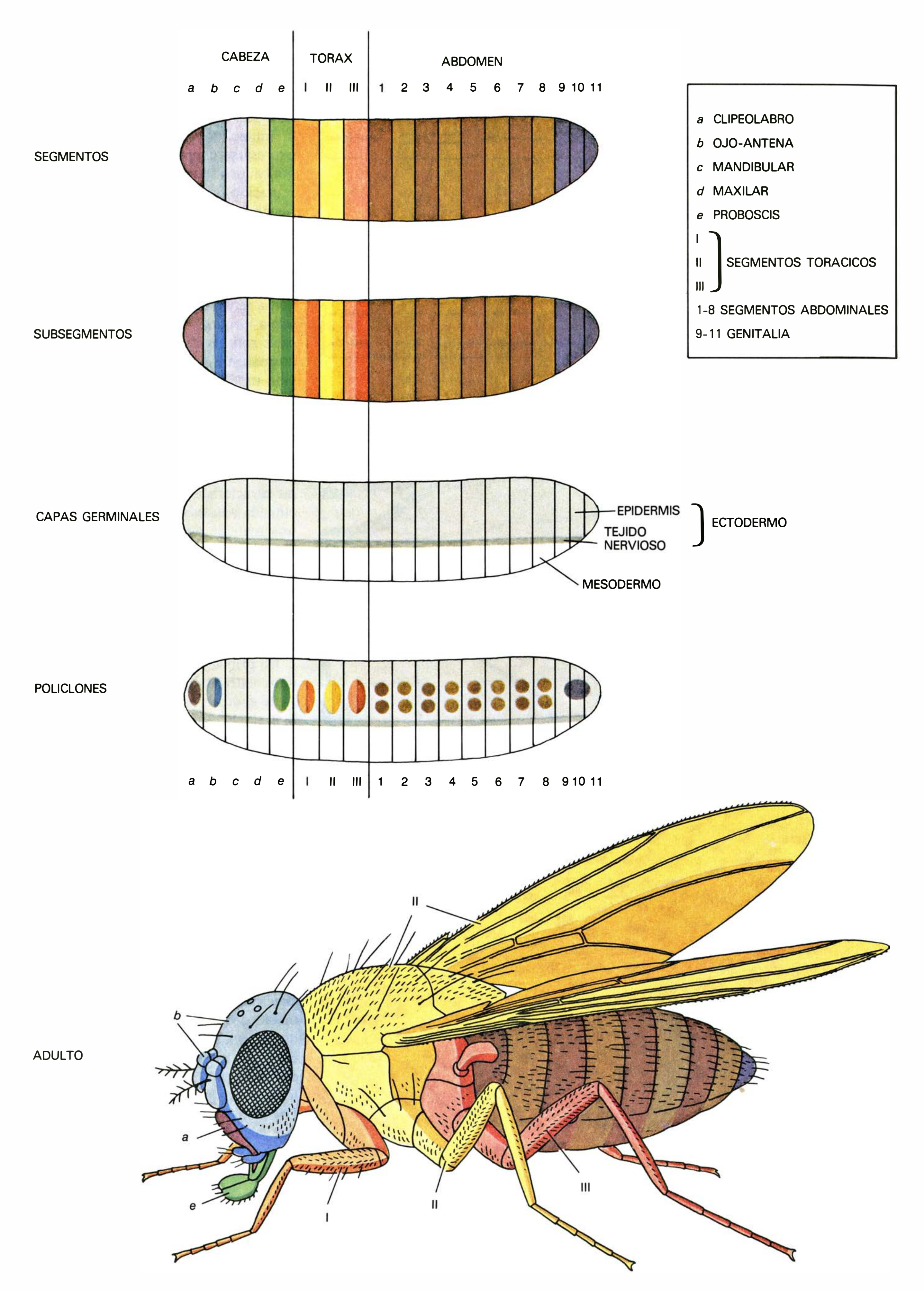
En muchos virus (por ejemplo, el virus del mosaico del tabaco) se conoce bastante bien el proceso mediante el cual la información genética se utiliza para construir la partícula viral. La secuencia de nucleótidos de los genes se traduce en una secuencia de aminoácidos. La diferente naturaleza de éstos, por sí misma, condiciona la forma tridimensional de las moléculas de proteínas de la cápside. Estas moléculas tienen una serie de sitios específicos para ensamblarse con otras moléculas iguales o distintas, constituyéndose así la cápside regular del virus. De esta manera, la forma final del virus se puede derivar directamente de la secuencia de nucleótidos.

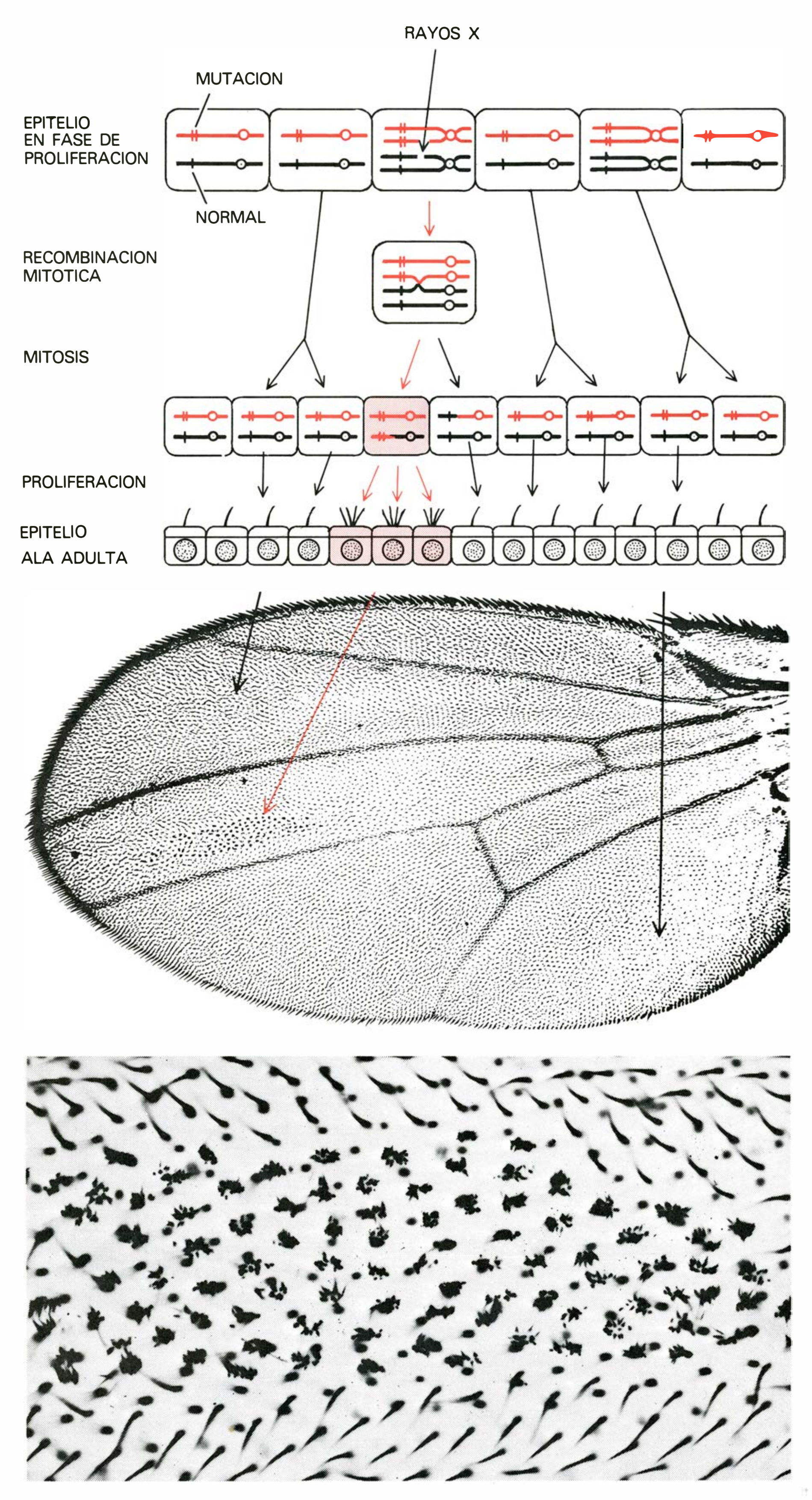
El proceso de morfogénesis en animales es, por supuesto, mucho más complicado. El huevo se subdivide en células que, aunque inicialmente son iguales, a medida que se dividen se van haciendo más y más diferentes. Las estructuras no se van formando por ensamblaje de partes ya terminadas, sino más bien por la progresiva elaboración de células y tejidos. Sin embargo, nosotros creemos que el desarrollo de virus y animales tienen en común el que tanto unos como otros parecen estar construidos de manera modular. Los virus están formados por subunidades idénticas mientras que los experimentos realizados recientemente, y que vamos a describir aquí, sugieren que los animales estén hechos de unidades, diferentes morfológicamente, pero fundamentalmente homólogas y que llamamos "compartimentos"; estos compartimentos serían en realidad variaciones de un tema básico y sus diferencias individuales estarían determinadas por la función de diferentes combinaciones de ciertos genes especiales.

Aunque los órganos pueden llegar a tener formas más o menos complejas debido a los plegamientos y engrosamientos que constituyen los tejidos, todos ellos se pueden reducir esencialmente a capas monocelulares. En consecuencia, el problema de la morfogénesis puede reducirse a un problema de dos dimensiones: ¿cómo controla la información genética la organización y estructura de las capas celulares? Para este tipo de pregunta, los insectos son un material particularmente favorable ya que manifiestan el problema en su forma más simple. Tienen una cutícula o exoesqueleto que es segregado por una capa monocelular llamada epidermis. Además, cada una de las células epidérmicas produce una pieza particular de la cutícula.

En la mosca del vinagre, *Drosophila melanogaster*, conocemos muchas mutaciones que cambian el color o la estructura de la cutícula de diversas maneras. Por ejemplo, en el ala de moscas normales cada una de las células produce un pequeño pelo llamado *tricoma*. Ahora bien, si las moscas llevan la mutación llamada *multiple wing hairs* cada célula da lugar a un grupo de 2-5 tricomas en vez de uno. A este tipo de mutaciones las llamamos mutaciones marcadoras. En el Instituto de Genética en Madrid del Consejo Superior de Investigaciones Científicas y en el Laboratorio de Biología Molecular del Medical Research Council en Cambridge, Inglaterra, hemos desarrollado una serie de técnicas (basadas en el trabajo pionero de Curt Stern, de la Universidad de California en Berkeley) para marcar células individuales en diversos momentos del desarrollo embrionario o larvario de *Drosophila* y poder estudiar su descendencia. Como resultado de un cruza-

COMPARTIMENTOS en el desarrollo de la mosca del vinagre *Drosophila melanogaster*. Se indican mediante diagramas esquemáticos del embrión y de la mosca adulta. El embrión, que en el blastodermo es un epitelio periférico de unas 5000 células, proviene de un mismo núcleo del huevo y se subdivide en segmentos destinados cada uno de ellos a formar un segmento adulto. Se sabe que varios segmentos de éstos se dividen en subsegmentos anterior (*sombreado claro*) y posterior (*sombreado oscuro*). Cada subsegmento se halla poblado por un policlón diferente. El epitelio se subdivide en capas germinales: ectodermo, que incluye la epidermis y el tejido nervioso, y mesodermo. Algunas células epidérmicas se seleccionan para formar los discos imaginales (*coloreados*) a partir de los cuales se desarrolla la cutícula adulta; el resto formará la epidermis o cutícula larvaria (los segmentos embrionarios *c* y *d* quizá no formen estructuras adultas). Los segmentos representan los compartimentos básicos del desarrollo de insectos según la hipótesis de los autores. Los policlones anterior y posterior construyen un grupo de compartimentos anteriores y posteriores y, a su vez, pueden subdividirse en otros policlones que producen compartimentos más reducidos. Ciertos genes "clave" están involucrados en este proceso.

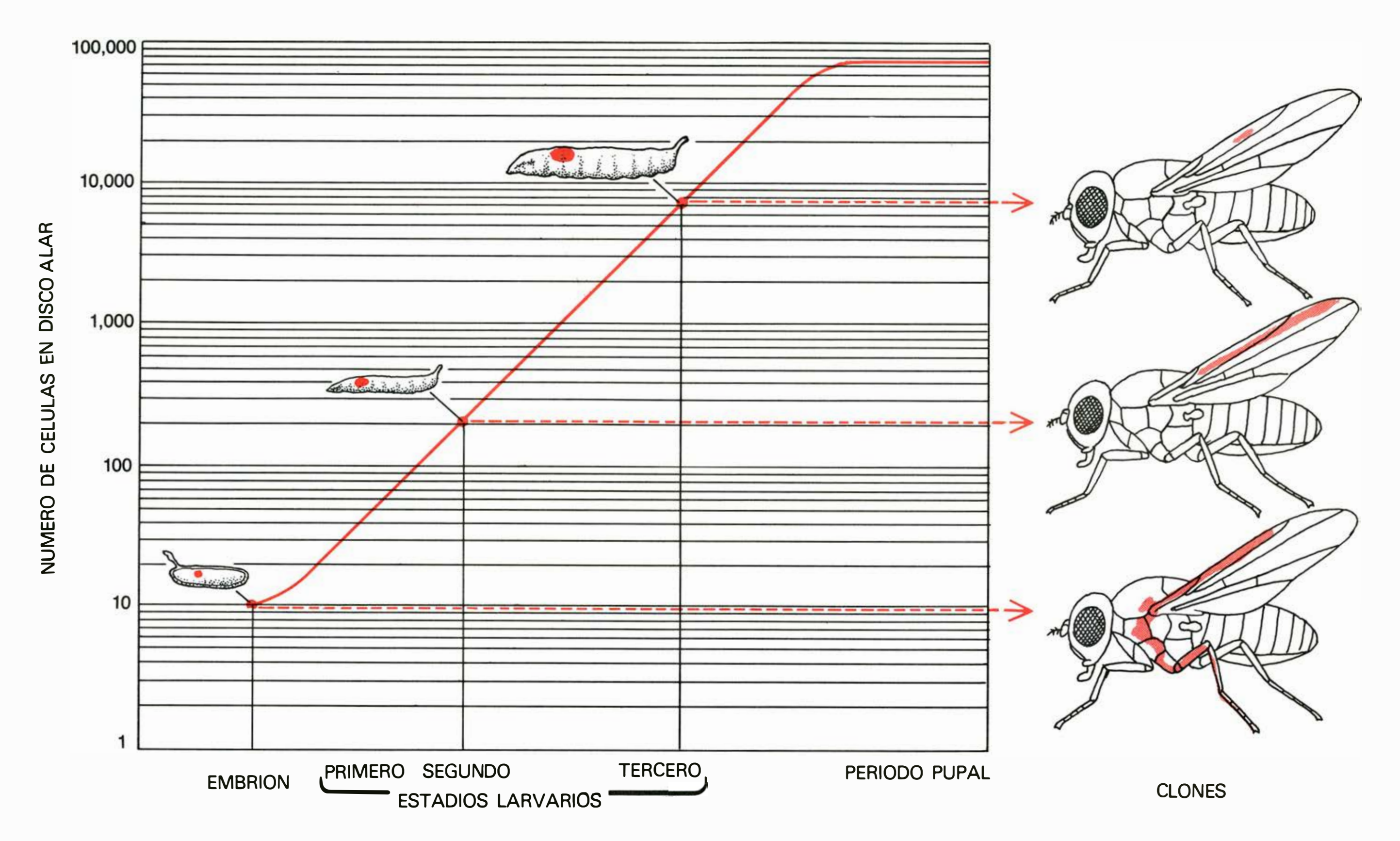

CABEZA
TORAX
ABDOMEN
a b c d e I II III 1 2 3 4 5 6 7 8 9 10 11
SEGMENTOS
SUBSEGMENTOS
CAPAS GERMINALES
EPIDERMIS
TEJIDO NERVIOSO
ECTODERMO
MESODERMO
POLICLONES
a b c d e I II III 1 2 3 4 5 6 7 8 9 10 11
a CLIPEOLABRO
b OJO-ANTENA
c MANDIBULAR
d MAXILAR
e PROBOSCIS
I
II
III
SEGMENTOS TORACICOS
1-8 SEGMENTOS ABDOMINALES
9-11 GENITALIA
ADULTO
b
a
e
I
II
II
III



LINAJE CELULAR estudiado en moscas que son heterozigóticas para una mutación recesiva que afecta a la cutícula. En este caso, una mutación (*m*) que determina grupos de pelos en vez de pelos aislados. Un cromosoma (*color*) del par homólogo lleva la mutación y el otro no; debido a esto, la mutación recesiva no se expresa normalmente. Si un embrión o larva se exponen a los rayos-X, la irradiación puede romper los cromosomas de tal forma que durante la mitosis o divisiones celulares, fragmentos de los cromosomas pueden "recombinar" y una célula se hace homozigótica para la mutación (los dos cromosomas del par homozigótico llevan la mutación). Esta célula y sus descendientes (*color claro*) producen pelos anormales. Una mancha de pelo anormal sobre el ala revela la existencia del clon mutante y define su extensión, como se ve en la micrografía (*centro*) y en su ampliación (*abajo*).

miento se obtienen moscas en las que cada célula es heterozigótica para una mutación marcadora, es decir, tiene un cromosoma que contiene la mutación y otro que contiene el gen normal. Al ser la mutación recesiva, solamente puede expresarse cuando la célula es homozigótica (los dos cromosomas llevan la mutación) y, por tanto, estas moscas tienen una apariencia normal. Ahora bien, cuando a un embrión o larva de estas características se le irradia con rayos X, se produce una serie de roturas cromosómicas que normalmente vuelven a sellarse recomponiendo el cromosoma original. Sin embargo, a veces estas roturas dan lugar a que trozos de cromosomas se recombinen (recombinación mitótica) y en la siguiente división celular, después de la irradiación, se produce una célula homozigótica para la mutación. Las sucesivas divisiones de esta célula originan un clon de células homozigóticas que, por tanto, expresan las características de la mutación. En la cutícula adulta aparecerá este clon como una mancha de tejido mutante rodeado de tejido normal. Se pueden irradiar cientos de embriones o larvas; luego, se comparan en las moscas adultas los clones producidos. De esta comparación podemos deducir el linaje celular de todas las estructuras cuticulares adultas.

En principio, uno puede imaginar dos tipos extremos de linajes celulares durante el desarrollo: fijo o indeterminado. Si el linaje celular fuera fijo, cada célula del embrión siempre acabaría por producir una parte específica del adulto y todos los embriones se desarrollarían de forma idéntica. Por el contrario, si el linaje fuera indeterminado, cualquier célula podría dar lugar a cualquier parte del adulto y las genealogías celulares variarían de un embrión a otro. El análisis de los linajes celulares del ala de *Drosophila* ha mostrado que ninguno de estos modelos extremos es correcto. Si se dibuja sobre un diagrama un número suficiente de clones encontrados en diferentes alas, vemos que los clones se superponen y sus bordes se cruzan. Esto quiere decir que, aunque la forma final de ala es siempre la misma, la contribución de las células que la hacen es diferente cada vez, esto es, el ala no se forma por un mecanismo de linaje fijo. Los clones marcados se pueden encontrar por toda la superficie del ala con igual frecuencia, lo que indica que el crecimiento es intercalar, es decir, el crecimiento se da en todas las zonas del ala y no en regiones específicas, como sucede por ejemplo en las plantas. Cuando se analiza el tamaño de los clones producidos en el mismo tiempo,



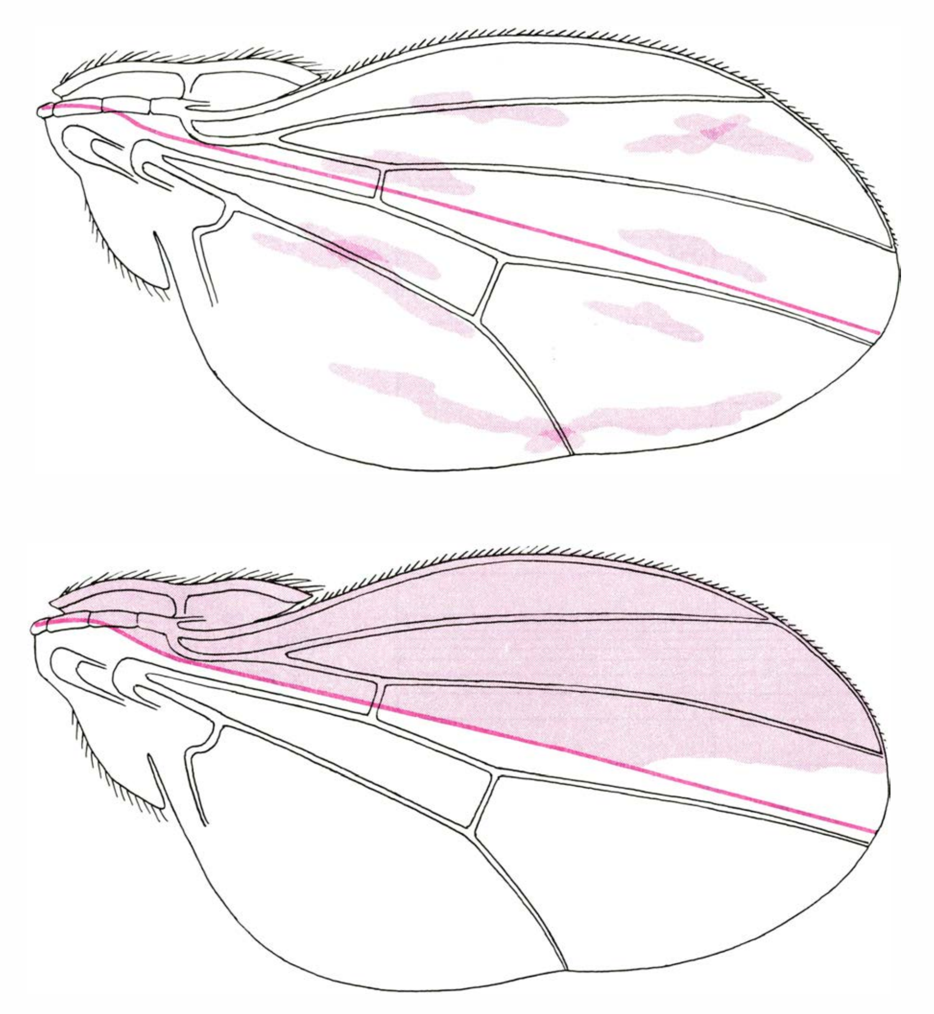
DEL ANALISIS DE LOS CLONES incluidos en sucesivos estadios de desarrollo se obtiene información sobre el crecimiento del ala y la pata. Si se irradia el pequeño grupo de células embrionarias (*color oscuro*) que dan lugar a la pata y el ala, el número de clones producidos es pequeño (la frecuencia de clones depende del número de células), pero son muy grandes pudiendo incluir el ala y la pata (*color claro*). En la larva, las células que producen las estructuras adultas forman unos órganos larvarios llamados discos imaginales. Es en esta fase cuando las células proliferan más rápidamente, de forma que la frecuencia de clones inducidos aumenta al haber más células, pero al mismo tiempo el tamaño de los clones es más pequeño.

se ve que pueden tener diferente número de células, lo que manifiesta que no hay un número fijo de divisiones celulares por clon.

Las observaciones anteriores eliminan la hipótesis del linaje celular fijo; no obstante, esto no quiere decir que haya una absoluta indeterminación de los linajes celulares. Como veremos a continuación hay ciertos límites en lo que la descendencia de una célula puede formar en la cutícula. Esto quedó muy claro en experimentos que llevamos a cabo hace algunos años en colaboración con nuestro colega Pedro Ripoll. En estos experimentos utilizamos moscas de unas estirpes especiales que, si bien son normales en cuanto a tamaño y forma, se desarrollan muy despacio debido a su lento ritmo de división celular. En estas moscas, por los procedimientos de recombinación mitótica descritos anteriormente, se producen clones marcados genéticamente, que además son capaces de dividirse a un ritmo normal, mucho más rápido que el de las células que los rodean. Es decir, estos clones disponen de mucho más tiempo para proliferar que las células vecinas. De hecho, en algunos casos, clones tempranos producidos por ejemplo en el ala tendrían tiempo para producir un número de células suficiente para formar varias alas. Sin embargo, estos clones quedan restringidos dentro de un ala de tamaño normal. Como se ve, este procedimiento permite explotar al máximo la capacidad de desarrollo de las células. Cuando estudiamos estos clones producidos en embriones tempranos encontramos que, a pesar de su gran tamaño, muestran ciertas restricciones topológicas en las estructuras de ala que producen. Existe una línea en la mitad del ala que nunca es cruzada por clones. Esta línea, respetada por todos los clones que alcanzan su posición, subdivide no sólo el ala sino todo el segundo segmento torácico (que incluye el "notum" y la segunda pata) en dos regiones, una anterior y otra posterior. El examen de clones pequeños producidos por la técnica convencional demuestra que estos clones también respetan la misma línea. Los clones grandes simplemente la dibujan mucho más claramente.

Este resultado significa que, en el momento en que estos clones se inician, al comienzo del desarrollo embrionario, el grupo de células que darán lugar al segundo segmento torácico están ya subdivididas en dos grupos, uno de ellos destinado a formar la región anterior y el otro la posterior del segmento. A estas regiones las llamamos *compartimentos*. Cada compartimento está formado por un conjunto de clones, que llamamos un *policlón* y que se desarrolla a partir de un pequeño grupo de células originarias. Cada policlón tiene un destino definido en cuanto al inventario de estructuras a las que dará lugar, aunque no así cada una de las células que lo forman, cuya contribución a la estructura final no está definida. (Como ya hemos visto, una célula que se divide rápidamente puede llegar a producir la mayor parte del compartimento mientras que las restantes células del policlón contribuyen mucho menos.)

Toda una serie de experimentos recientes han demostrado que también otras partes de la mosca se construyen mediante compartimentos que han sido identificados hasta ahora en los tres segmentos torácicos, en la cabeza, el abdomen y las estructuras genitales externas. Algunos de estos compartimentos están delimitados por discontinuidades de la cutícula, como en el caso de los bordes de los segmentos que forman la mosca adulta, pero esto no es siempre así. Por ejemplo, no hay ninguna característica morfológica que defina la línea de demarcación anteroposte-



***DROSOPHILA* no tiene un linaje celular fijo. Si se dibujan una serie de clones típicos producidos en alas diferentes sobre un diagrama, se ve que los clones se superponen unos con otros y sus bordes se cruzan; esto quiere decir que los descendientes de cada célula marcada producen diferentes partes del adulto en cada caso (*arriba*). Sin embargo, el linaje no es completamente indeterminado; cuando se producen clones que proliferan rápidamente, nunca dan lugar más que a la mitad del ala (*abajo*), bien la mitad anterior (como en el dibujo) o la mitad posterior. Cuando se estudian con detalle los clones celulares más pequeños (*arriba*), se observa que también éstos respetan la misma línea de demarcación.**

rior en el ala. Esta línea es invisible a menos que se utilicen las técnicas de marcaje genético.

No obstante, aunque las pruebas de la existencia de compartimentos en la cutícula son bastante sólidas, sabemos todavía muy poco de los linajes celulares de los órganos internos de la mosca. La razón de esta ignorancia se debe fundamentalmente a la falta de marcadores celulares apropiados. No sabemos, por ejemplo, si existen compartimentos en el intestino, los músculos o el sistema nervioso central.

Este tipo de experimento de marcaje celular ha establecido dos hechos importantes acerca de la construcción de los compartimentos; el primero de ellos es que los policlones se establecen progresivamente; los nuevos policlones se forman por subdivisión de los preexistentes; ya hemos visto que el segundo segmento torácico se forma a partir de los policlones anterior y posterior, los cuales se establecen en el embrión temprano. Poco tiempo después se da una subdivisión que separa los policlones que forman el ala de los que forman la pata y que afecta tanto al compartimento anterior como al posterior. Más tarde, los que van a formar el ala se subdividen en policlones hijos que van a originar la parte dorsal y ventral del ala. La historia de esta progresión no se conoce en su totalidad; no obstante, a título de ejemplo hemos elaborado un posible árbol de decisiones en que cada bifurcación representa un punto donde el policlón se subdivide en dos nuevos.

El segundo hecho establecido es que la misma decisión se puede tomar en diferentes segmentos. La subdivisión anterior-posterior, por ejemplo, tiene lugar en los tres diferentes segmentos torácicos y en la cabeza. Asimismo, la subdivisión dorsoventral tiene lugar en al menos dos segmentos torácicos. Cada rama final del árbol de decisiones está especificada por un código binario que recoge todas las decisiones tomadas durante el curso del desarrollo. Dentro del segundo segmento torácico, por ejemplo, el código para anterior-pata-dorsal y el código para anterior-ala-dorsal sólo difieren en una decisión, la que separa las células de ala y pata; el resto es común.

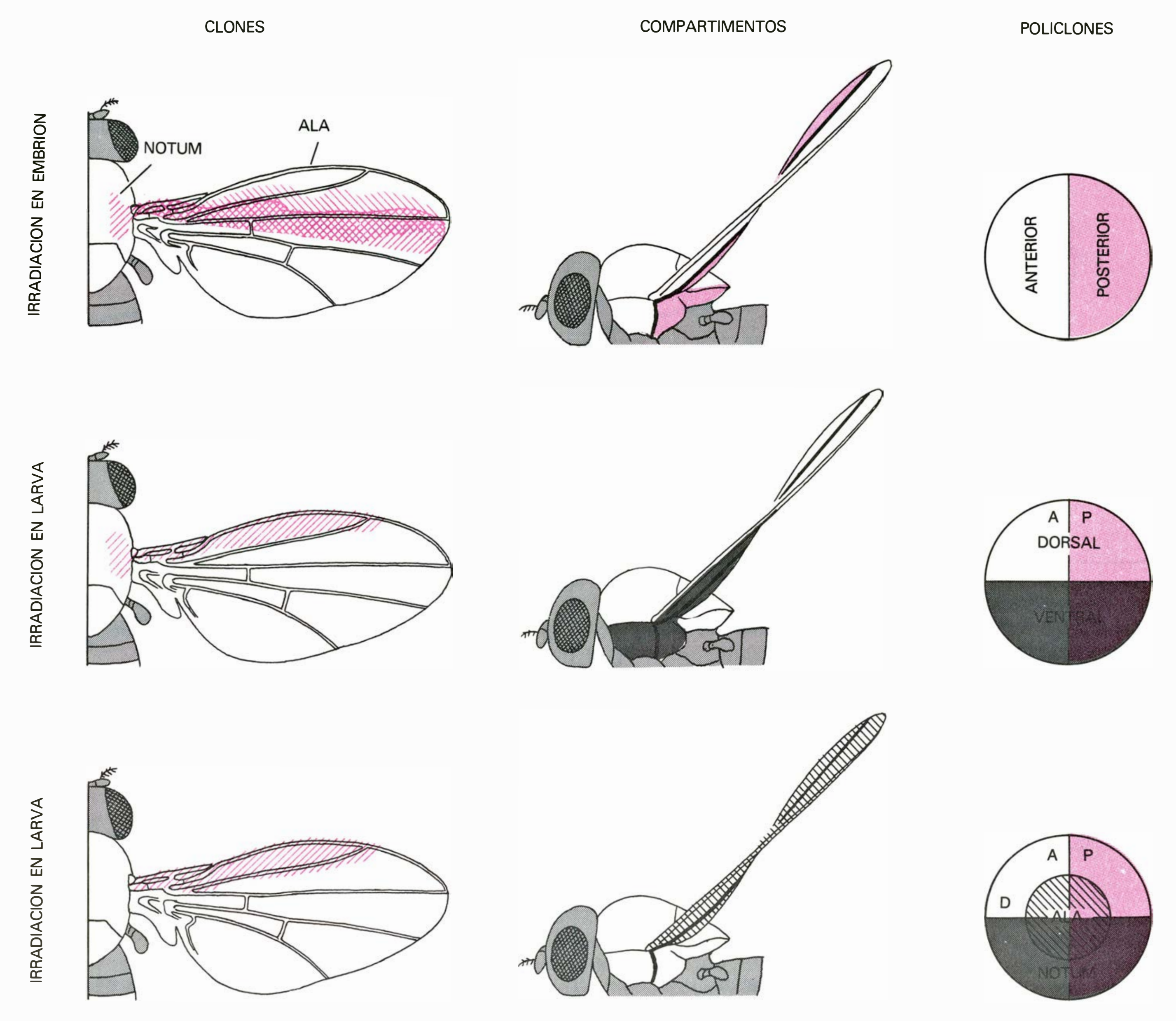
¿Cómo se toman estas decisiones y cuál es su naturaleza? Sabemos muy poco todavía acerca de cómo se subdividen estos policlones, aunque desde luego sabemos que la decisión tiene una base topológica de forma que los policlones se subdividen en grupos de células de acuerdo con su posición original. También sabemos que el grupo inicial de células de un policlón no están originadas clonalmente; es decir, no existe una célula madre que dé lugar exclusivamente a las células de un policlón. Otro hecho importante demostrado mediante análisis clonal es que este tipo de decisiones que define compartimentos son decisiones definitivas (al menos durante el curso normal de desarrollo) y permanentes; toda la descendencia de una célula de un policlón contribuirá solamente a un compartimento determinado.

Vamos a considerar ahora el papel que juegan ciertos genes en el desarrollo de los compartimentos. En *Drosophila* se conocen desde hace bastante tiempo unas mutaciones llamadas "homeóticas", que transforman ciertas regiones específicas del cuerpo de la mosca en otras. Cada una de estas mutaciones da lugar a una transformación en particular y se localiza en sitios específicos del cromosoma, lo cual quiere decir que son mutaciones en diferentes genes. A título de ejemplo, *Antennapedia* transforma la antena de la mosca en pata, mientras que otra llamada *Opthalmoptera* transforma ojo en ala. Hay una serie de mutaciones homeóticas ligadas genéticamente que forman la serie *bithorax*, que ha sido estudiada durante muchos años por Edward B. Lewis, del Instituto de Tecnología de California. Las mutaciones en el gen *bithorax* transforma la parte anterior del tercer segmento torácico en la parte anterior del segundo segmento torácico. La transformación afecta lo mismo a la parte dorsal que a la ventral del segmento, pero solamente se extiende hasta la línea que separa el compartimento anterior del posterior. La mutación *postbithorax* transforma la parte posterior del tercer segmento en la posterior del segundo segmento torácico; de nuevo,

el límite de los compartimentos anterior y posterior. Las mutaciones en el gen *engrailed* afectan a las regiones posteriores de ala y patas, transformándolas en las correspondientes anteriores del mismo segmento y, una vez más, el límite de la transformación coincide con el límite de los compartimentos anteriores y posteriores.

Las características (fenotipos) de las mutaciones *bithorax, postbithorax* y *engrailed* sugieren una conclusión general: Los compartimentos son los dominios de función de al menos algunos genes homeotípicos, o dicho de otra forma, los policlones son los grupos de células sobre los que operan los genes homeóticos. Existen pruebas hoy día de que posiblemente se necesite la función del gen *engrailed* en todos aquellos casos en que se requiere una propiedad posterior, ya que la mutación (que representa la falta de función normal del gen) afecta a los compartimentos posteriores, no sólo del ala, sino también de las tres patas y el segmento del ojo-antena, sin que en ningún caso afecte a los compartimentos anteriores.

Estas observaciones nos han llevado a la idea de que el árbol de decisiones binarias describe en realidad la adquisición progresiva de una instrucción o "cifra genética" para cada uno de los compartimentos que aparecen sucesivametne. En cada punto de decisión hay un gen clave que aparece en uno solo de dos estadios posibles"on" u "off", y que es el mismo gen que opera en decisiones homólogas tomadas en diferentes segmentos o compartimentos. Por ejemplo, la función del gen *engrailed* lleva a cabo la decisión "posterior". Este gen tiene que estar activo en todos los compartimentos posteriores y es inactivo en todos los compartimentos anteriores; por esta razón, cuando el gen no funciona o no lo hace de forma normal (como en el caso de la mutación *engrai-*



SECUENCIA TEMPORAL de la formación de los compartimentos del ala y del notum. Al irradiar embriones (*serie superior*), se producen clones grandes (*izquierda*) que se extienden por el compartimento (*centro*) anterior (*blanco*) o por el posterior (*coloreado*) pero nunca a ambos a la vez, ya que en el momento de la irradiación el disco estaba subdividido en policlones anterior y posterior (*derecha*). Los clones que se producen en el período larvario (*serie central*) están confinados bien a la parte dorsal (*blanco*) o a la ventral (*gris oscuro*) de los compartimentos anteriores y posteriores (*centro*); en este momento, el disco está ya subdividido en cuatro policlones (*derecha*). Más o menos al mismo tiempo, los clones aparecen confinados (*serie inferior*) al notum (*blanco*) del ala (*punteado*). Los cuatro policlones preexistentes se han vuelto a subdividir cada uno de ellos en dos. Cada uno de los ocho policlones resultantes (*abajo, derecha*) posee una combinación diferente de tres propiedades a saber: anterior o posterior, dorsal o ventral, ala o notum.

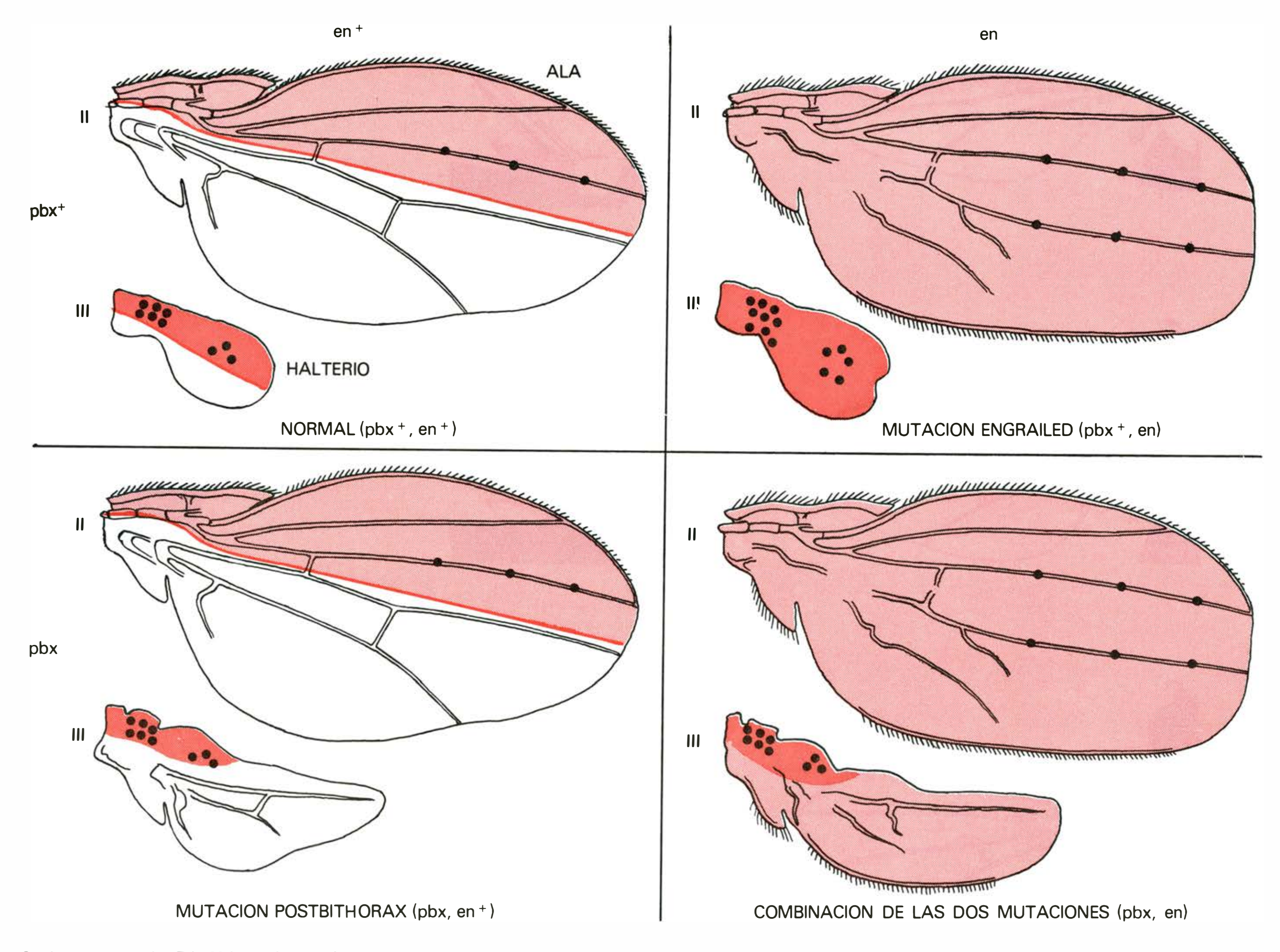
led) los compartimentos posteriores se transforman en anteriores.

Esta hipótesis predice claramente que cada mutación homeotípica cambia una sola propiedad, y eso es precisamente lo que se observa. Cuando, por ejemplo, la mutación *engrailed* transforma posterior en anterior, las decisiones dorsal-ventral y notum-ala no quedan afectadas. Además, implica también que las mutaciones homeóticas producen fenotipos aditivos como se observa en la práctica en el caso de los mutantes dobles.

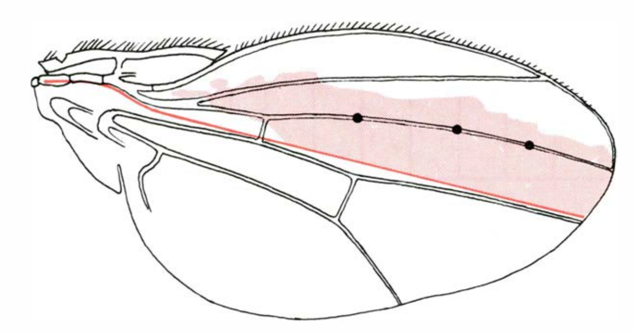
En nuestra opinión, la combinatoria de genes homeóticos activos o inactivos proporciona una instrucción o señas genéticas que se mantiene permanente en cada célula y que se transmite a las células hijas. Un resultado que apoya esta opinión es que se pueden cambiar cifras genéticas simplemente cambiando el estado funcional de uno de los genes clave. Consideremos por ejemplo *engrailed.*

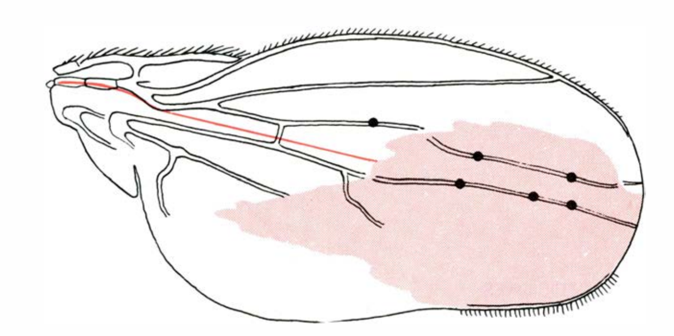
Por medio de la técnica de recombinación mitótica descrita anteriormente, podemos producir en un ala normal un clon de células que son homozigóticas para la mutación *engrailed*. En la parte anterior del ala este cambio no produce ningún efecto, porque en nuestra hipótesis no hay función del gen *engrailed* en este compartimento. Sin embargo, sí tiene efecto en la región posterior, donde la actividad del gen normal de *engrailed* es requerida (para mantener la condición posterior) y, por tanto, su falta de función en una célula cambia las cifras genéticas de ella y de sus descendientes. El resultado es que el clon de células homozigóticas desarrolla estructuras anteriores. Este experimento indica que lo que las células hacen en el desarrollo depende muy fundamentalmente de sus cifras genéticas.

El comportamiento de estos clones *engrailed* no depende sólo de su cifra genética, sino también de su posición dentro del comportamiento. El borde anterior del ala contiene unas cerdas cortas y gruesas que no aparecen en el borde posterior. Cuando se produce un clon *engailed* que incluye parte del borde posterior del ala, las células pertenecientes al clon y que están situadas en el borde diferencian este tipo de cerdas gruesas anteriores. Las células del clon que no están en el borde no las producen, ni por supuesto tampoco lo hacen las células exteriores al clon, aunque estén en el borde. Esto indica que el patrón morfológico final depende tanto de la posición de las células en el órgano como de su cifra. Las célula mutantes para *engrailed* y las normales que las rodean "saben", de alguna forma, su posición en el compartimento y responden interpretando su posición diferente-



LAS MUTACIONES HOMEOTICAS son aditivas, como se demuestra por los efectos (que se han esquematizado para mayor claridad) de las mutaciones *engrailed* (*en*) y *postbithorax* (*pbx*) en el segundo (*II*) y tercero (*III*) segmentos torácicos. Normalmente (*arriba, izquierda*), el segundo segmento torácico forma un ala con un compartimento anterior (*color*) y otro posterior (*blanco*); el tercer segmento torácico forma un apéndice llamado halterio, que se halla subdividido también en compartimentos anterior y posterior. La mutación *engrailed* transforma los segmentos posteriores en anteriores (*arriba, derecha*); *postbithorax* transforma el compartimento posterior del tercer segmento en el compartimento posterior del segundo segmento; esto es, transformando parcialmente el halterio en ala (*abajo, izquierda*). En el caso de la doble combinación mutante, *postbithorax* produce un apéndice mitad halterio y mitad ala, mientras que *engrailed* transforma el compartimento posterior de este apéndice mixto en compartimento anterior de ala.





EL EFECTO DE *ENGRAILED* se puede estudiar en clones mutantes producidos en alas normales. Las células de estos clones *engrailed* están marcadas por una segunda mutación que permite distinguir las células mutantes, aun cuando *engrailed* no tuviera efecto. Los clones *engrailed* (*color*) del compartimento anterior se desarrollan con normalidad y respetan la línea de demarcación antero-posterior (*izquierda*). Los clones *engrailed* del compartimento posterior, por el contrario, manifiestan las estructuras y el patrón morfológico típicos del compartimento anterior. Este compartimento lo muestran exclusivamente las células del clon. El tipo de estructuras anteriores producidas por el clon depende de su posición; las típicas cerdas del margen anterior del ala sólo son producidas por las células del clon en el margen posterior. Estos clones no respetan la línea divisoria antero-posterior. Ello obedece a que, a las células posteriores mutantes, les falta una "marca" en su membraba que impide que se mezclen con células anteriores.

mente. Parece ser que la información posicional es la misma en el compartimento anterior y en el posterior, pero la interpretación de esa información es diferente (de una forma similar a lo que sucede cuando se cubren con una misma cuadrícula mapas de ciudades distintas: la misma referencia posicional designa sitios muy diferentes, aunque se hallen en el mismo cuadro sobre el mapa).

Otras características de los clones mutantes para *engrailed* es que no se respetan la línea de demarcación entre los compartimentos anterior y posterior (aunque desde luego respeten las líneas de demarcación de otros compartimentos). Las células mutantes que se originan en el compartimento posteior pueden invadir el compartimento anterior, algo que las células normales jamás hacen. Nosotros interpretamos este resultado indicando que las células posteriores tienen una "marca" en la superficie distinta de las anteriores, de forma que el contacto entre las células que llevan la marca y las que no la llevan es mínimo. Hemos sugerido que esta marca resulta de la función normal del gen *engrailed*. Así, mientras que las células normales anteriores y posteriores no pueden mezclarse, los clones mutante sí pueden, no respetan la línea de demarcación y se mezclan con células anteriores. Cada uno de los policlones tiene una combinatoria de marcas única que le confieren unas propiedades de reconocimiento celular específico (los experimentos llevados a cabo por Erns Hadorn y colaboradores en la Universidad de Zürich han establecido que células marcadas provenientes de distintas partes de la mosca que han sido artificialmente disociadas son capaces de reagruparse, juntándose cada una de ellas con las de su mismo origen y segregándose de los de origen distinto. En nuestra opinión, el propósito de estas marcas es doble: mantener diferentes compartimentos separados sin que haya mezcla, y, además, ayudar en el mecanismo de ensamblaje para que los diferentes compartimentos encajen perfectamente.

¿Cómo funcionan estos genes homeóticos? Hasta ahora sólo hay unos pocos indicios; ni siquiera se conoce la naturaleza de sus productos génicos. Se podría esperar que cada uno de estos genes clave controlasen un equipo específico de genes subsidiarios. Si éste fuera el caso, la mayor parte de los genes de *Drosophila* funcionarían en un compartimento o en un subconjunto de compartimentos, y el hecho es que la mayoría de las mutaciones no son exclusivas de compartimentos. La mutación *white*, por ejemplo, es conocida primeramente por su efecto en el ojo, pero además afecta a los pigmentos en los tubos excretores, los testículos y otros órganos; otras mutaciones que cambian los pelos o cerdas de la cutícula afectan a todo el cuerpo de la mosca. Lo más probable es que los genes homeóticos operen básicamente sobre el mismo grupo de genes subsidiarios.

En resumen, el modelo que hemos desarrollado pretende describir la estrategia genética que se sigue para construir un insecto (un compartimento hecho por un grupo particular de células): es específica mediante una cifra genética, que resulta de una serie de decisiones (en código binario) de activación o inactivación de una serie de genes clave. Las diferentes cifras genéticas se establecen a partir de un árbol de decisiones en el que, en cada bifurcación, un gen se activa en una rama mientras que es inactivado en la otra. El estado de función del gen se mantiene de la misma forma en todas las células hijas, de modo que en cada bifurcación se incorpora un nuevo elemento a la cifra genética. Así pues, cada célula adulta contiene la historia de las decisiones tomadas por sus células progenitoras. Esta estrategia da lugar a que compartimentos de diferentes posiciones tengan un cifra genética distinta.

En este modelo, un pequeño número de elementos de control genético son suficientes para especificar un número considerable de compartimentos diferentes. El modelo permite además imaginar cómo la evolución ha ido modificando los segmentos de los insectos a partir de los más primitivos, en los que todos los segmentos son idénticos. Tan pronto como un elemento genético se hace cargo del desarrollo de un segmento individual, este segmento puede evolucionar de forma independiente del resto. Mientras que el número de segmentos y la lógica genética general del proceso de control permanezce estable durante el curso de la evolución, la morfología de cada segmento individual puede cambiar rápidamente. Las mutaciones homeóticas conocidas en otros insectos apoyan esta idea. Se conocen transformaciones homeóticas en mosquitos y escarabajos similares a las halladas en *Drosophila*, aunque las estructuras cuticulares involucradas en estas transformaciones son diferentes.

Hay que hacer notar que incluso en el caso de que estas hipótesis sean correctas, hay muchos problemas que permanecen sin resolver. Por ejemplo, no comprendemos en absoluto cómo se establecen los patrones morfológicos en cada segmento. La información posicional parece ser crucial para las decisiones genéticas que se toman dentro de los policlones y también para determinar la

CABEZA TORAX ABDOMEN

EMBRION

OJO-ANTENA PROBOSCIS I II III

DECISIONES

ANTERIOR/POSTERIOR

LARVA/ADULTO

VENTRAL/DORSAL EN EL SEGMENTO

DORSAL/VENTRAL EN EL DISCO IMAGINAL

TRONCO/APENDICE

CIFRA GENETICA BINARIA

0 1 1 0 1

ESPECIFICA EL COMPARTIMENTO ANTERIOR, DORSAL DEL ALA ADULTA

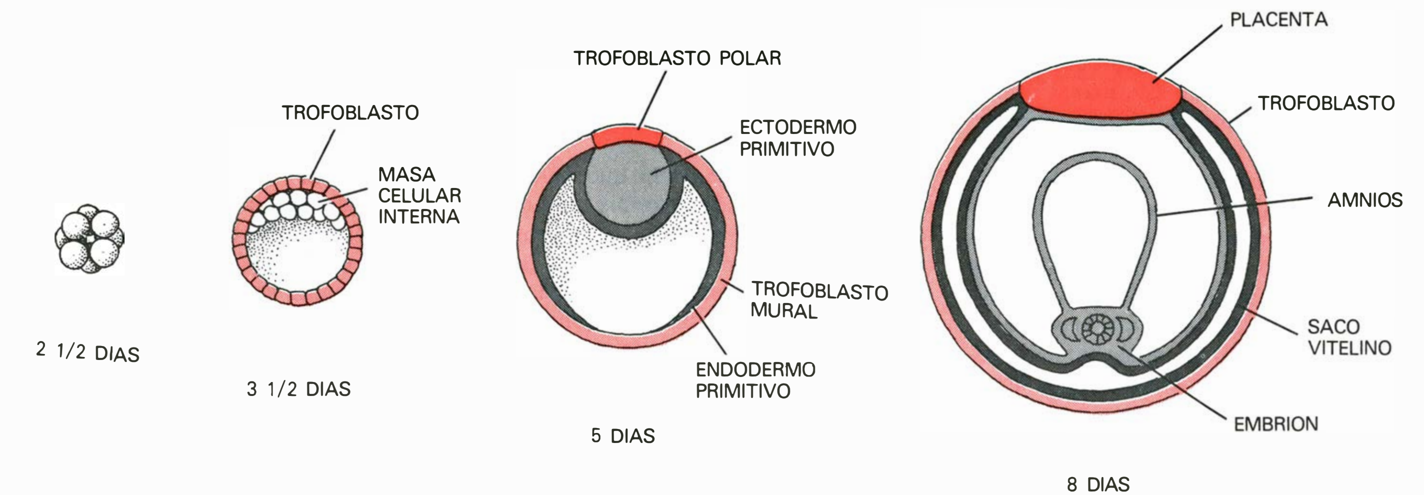
1 1 0

ESPECIFICA EL COMPARTIMENTO POSTERIOR DE LA PATA

ETAPAS SUCESIVAS de la subdivisión de los policlones. Pueden representarse como un árbol genealógico de decisiones "on" (*1*) y "off" (*0*) de ciertos genes clave. El diagrama es especulativo y muestra lo que, en opinión de los autores, podría ser el proceso. Se toman decisiones homólogas en segmentos diferentes. La decisión antero-posterior tiene lugar en cinco segmentos al menos. Se han identificado otras decisiones, tal como se indica en la figura. La mutación *engrailed* afecta a la separación de células anteriores y posteriores; otros genes clave controlarían las restantes decisiones, pero éstos siguen sin conocerse. La serie de decisiones da lugar a una "cifra genética" según un código binario que especifica para cada compartimento.

forma y tamaño de los compartimentos individuales, y, aunque la importancia de la información posicional en el desarrollo de los animales ha sido establecida claramente [véase "La formación de modelos en el desarrollo biológico", por Lewis Wolpert, INVESTIGACION Y CIENCIA, diciembre, 1978] seguimos sabiendo muy poco de los mecanismos que hacen que las células se comporten de forma adecuada a su posición.

¿Hay algún tipo de pruebas de la existencia de compartimentos en otros animales, además de los insectos? Ciertamente hay muchos animales que muestran una estructura a base de unidades repetidas, desde gusanos hasta vertebrados. En el caso del ratón, se sabe lo suficiente sobre linajes celulares como para preguntarse si tienen compartimentos los ratones. A la hora de responder hay que tener en cuenta una serie de criterios diagnósticos que definen la existencia de compartimentos: (1) pequeños



DESARROLLO DE LOS MAMIFEROS. Cabe que también en éstos ocurra un proceso de compartimentalización. El huevo fertilizado se divide hasta formar una pelota de células, la mórula (*izquierda*). En el estadio de blastocito (3 1/2 días), las células están ya subdivididas en dos policlones: la masa celular interna y el trofoblasto. A los cinco días se distinguen cuatro policlones: trofoblasto polar, trofoblasto mural, endodermo primitivo y ectodermo primitivo; cada una de estas cuatro capas celulares da lugar a estructuras específicas y, a buen seguro, se subdividen en otros policlones.

grupos de células (del orden e las decenas) dan lugar a grandes y defindas regiones en el adulto; (2) la decisión de qué células van a formar parte del grupo inicial de células fundadoras de un policlón se toma en base a la posición de las células y no según su linaje; (3) se establece un árbol de decisiones binaris durante el desarrollo; (4) células de diversos policlones manifiestan diferentes propiedades de reconocimiento celular; y (5) mutaciones homeóticas pueden cambiar el desarrollo de un policlón y esta forma cambiar un órgano por otro.

Se están acumulando pruebas, particularmente a partir del trabajo de Richard K. Gardner y sus colegas de la Universidad de Oxford, de que en el desarrollo temprano del ratón se verifican los cuatro primeros criterios diagnósticos que definen los compartimentos. Los métodos de análisis del grupo de Oxford son diferentes de los nuestros y requieren una mayor intervención experimental. Básicamente, su método consiste en mezclar dos embriones diferentes o en transplantar células de un embrión a otro. Después, siguen el destino de as células en el embrión en mosaico con la ayuda de anticuerpos fluorescentes que se ligan sólo a las células de uno de los donantes.

En ratones (como en otros mamíferos) el huevo fertilizado se empieza por dividir varias veces hasta formar la mórula, de forma más o menos esférica y que contiene de 16 a 32 células. A medida que la mórula se transforma en blastocisto, sus células se van a subdividir en dos grupos distintos que darán lugar a dos regiones diferentes: las células de la parte interna van a constituir lo que se llama la masa celular interna y el resto de las células forman una capa periférica que se llama el trofoblasto. Esta observación satisface el primero de los criterios mencionados anteriormente. Están en marcha experimentos para saber cuándo las células se definen de manera irreversible par formar una de estas dos estructuras. Lo que a nosotros nos interesa resaltar es que durante el desarrollo normal de un tipo celular, éste se subdivide en dos, y tal decisión parece tomarse según una base topológica, satisfaciendo por tanto el segundo criterio. Subsecuentemente, la masa interna de células se divide para formar dos policlones, satisfaciendo el tercer criterio. De nuevo la decisión se toma según posición, ya que las células cercanas al borde de la masa interna se convierten en el endodermo primitivo (que dará lugar a determinadas partes de las membranas extraembrionarias), mientras que el resto de las células forman el ectodermo primitivo (que da lugar al feto y otras estructuras).

Cuando se transplantan células aisladas de la masa celular interna de un blastocisto a otro de diferente genotipo (de un blastocisto de rata a otro de ratón, por ejemplo) y se sigue el destino de sus descendientes, se encuentra que éstos aparecen exclusivamente formando parte de las estructuras que derivan de una de las dos regiones de la masa celular interna, bien el endodermo primitivo o el ectodermo primitivo. Este resultado indica que hay dos (y sólo dos) tipos celulares distintos en la masa celular interna. Significa, además, que cada tipo celular donante posee la capacidad de integrarse en la región apropiada del embrión huésped, lo que sugiere que los dos tipos celulares tienen propiedades de reconocimiento específico. Esta última observación satisfaría el cuarto criterio.

A pesar de estas pruebas, que sugieren una rígida determinación de tipos celulares discretos en el desarrollo temprano del ratón, una serie de estudios realizados con ratones mosaico´que tienen diferentes coloraciones en la piel y otros marcadores demuestran que las células embrionarias se mezclan en una extensión considerable antes del nacimiento del ratón. El resultado final es bastante diferente de lo que sucede en la cutícula de *Drosophila*, en la cual las células que constituyen un clon generalmente permanecen juntas. Esta distinción entre el desarrollo de los insectos y los mamíferos podría no ser cierta en todos los tejidos ni en todo momento. En el ratón los clones muestran menos mezcla de embrión temprano que en el animal recién nacido y, en insectos, hay mucha más mezcla en la cutícula del abdomen y en los túbulos excretores que en la cutícula cefálica o el tórax. Es posible que dentro de un policlón el hecho de que las células se mezclen no tenga mayor importancia al ser todas ellas genéticamente equivalentes.

En el ratón, como en *Drosophila*, el proceso mediante el cual los policlones se subdividen en policlones hijos se halla todavía en fase de investigación. Hasta el momento, no se reconocen mutaciones homeotípicas en el ratón. Sin embargo, hemos observado que los ratones y las moscas tienen una serie de características de desarrollo en común, que permiten preguntarnos si acaso los insetos, el ratón y el hombre tienen un tipo de desarrollo basado en una estrategia genética parecida y que se expresa en la formación de compartimentos.

El modelo en bolsa del confinamiento de los quarks

Los quarks parecen ser partículas reales, a pesar de que todavía no se haya logrado su aislamiento individual. Una hipótesis que explica por qué no se han visto es que están confinados en bolsas, como las burbujas en los líquidos

Kenneth A. Johnson

Las partículas subatómicas se clasifican en dos grandes categorías: los leptones (el electrón, el muon, la partícula tau y los neutrinos asociados con cada uno de ellos), que no interaccionan fuertemente y parecen no tener partes constituyentes, y los hadrones (entre los que se cuentan el protón, el neutrón y el pion), que interaccionan fuertemente y muestran signos de una estructura interna. La fuerza que mantiene a los neutrones y protones juntos en el núcleo atómico es una manifestación de la interacción fuerte entre hadrones. Esta interacción constituye una de las cuatro fuerzas de la física, siendo las restantes la fuerza electromagnética, la fuerza débil y la gravitación.

En las últimas décadas se han identificado más de 200 clases de hadrones. También se les ha dividido en dos categorías, en función de la manera como se desintegran. Los bariones son las partículas que se desintegran en su última fase en el protón, y los mesones son las partículas que se desintegran completamente en leptones y fotones (quanta de la energía electromagnética) o en pares protón-antiprotón.

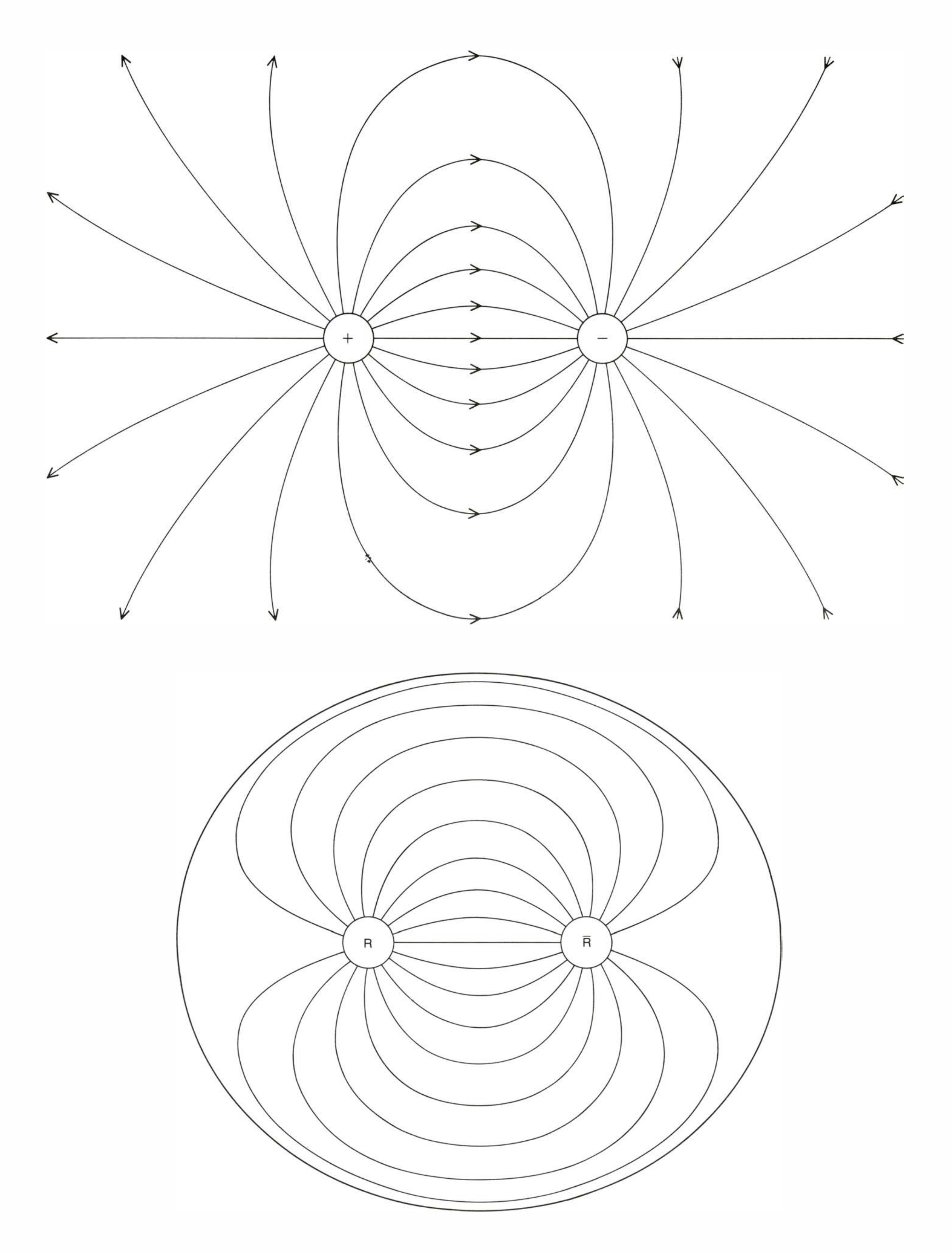
Todos los hadrones parecen ser combinaciones de unos constituyentes denominados quarks. El modelo de quarks representa una impresionante simplificación de la naturaleza, porque en la actualidad con sólo cinco "sabores", o clases, de quarks (y en la mayoría de casos con sólo tres sabores) se puede explicar las propiedades de la multitud de hadrones. La interacción fuerte mantiene los quarks juntos formando bariones y mesones. En el modelo de quarks, los bariones constan de tres quarks y los mesones están formados por un par quark-antiquark. De los quarks, cuyos sabores se llaman fantasiosamente "arriba" ("up"), "abajo" ("down"), "extraño" ("strange"), "encantado" ("charmed") y "fondo" ("bottom") nos intriga que no se haya observado ninguno de ellos aislado, a pesar de los muchos intentos realizados para su detección. Al objeto de buscar una explicación a este fenómeno, el autor y un grupo de colaboradores (Allen Chodos, Robert L. Jaffe, Charles B. Thorn y Victor F. Weisskopf) han desarrollado, en el Instituto de Tecnología de Massachusetts, un modelo teórico en el que los quarks que forman un hadrón están atrapados en lo que puede llamarse una bolsa o burbuja. El modelo sirve para clarificar muchos aspectos de la interacción fuerte entre los quarks.

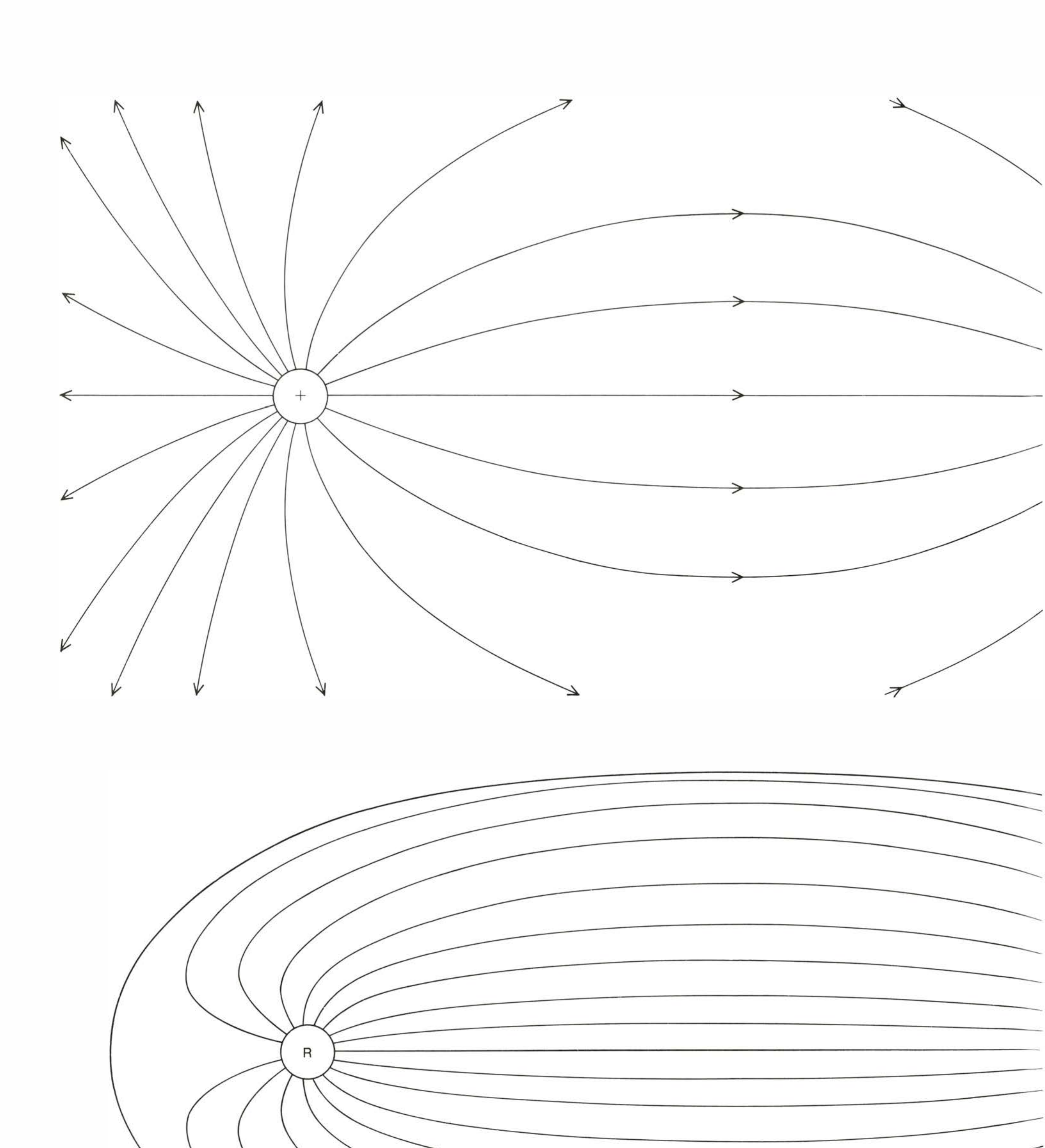
Cuando Murray Gell-Mann y George Zweig, del Instituto de Tecnología de California, introdujeron el modelo de quarks en 1963, no se los suponía universalmente como entidades reales que pudieran encontrarse en el laboratorio. Se consideró la posibilidad de que fueran puras construcciones teóricas que ponían orden en la engorrosa colección de hadrones. Corría la opinión de que los quarks eran objetos peculiares porque tenían sólo una fracción (2/3 o −1/3) de la carga eléctrica del protón, mientras que todas las otras entidades físicas conocidas tenían un múltiplo entero de dicha carga. Además, parecían violar el principio de exclusión, el equivalente mecánico-cuántico de la noción intuitiva de que dos cosas no pueden estar en el mismo sitio en el mismo instante.

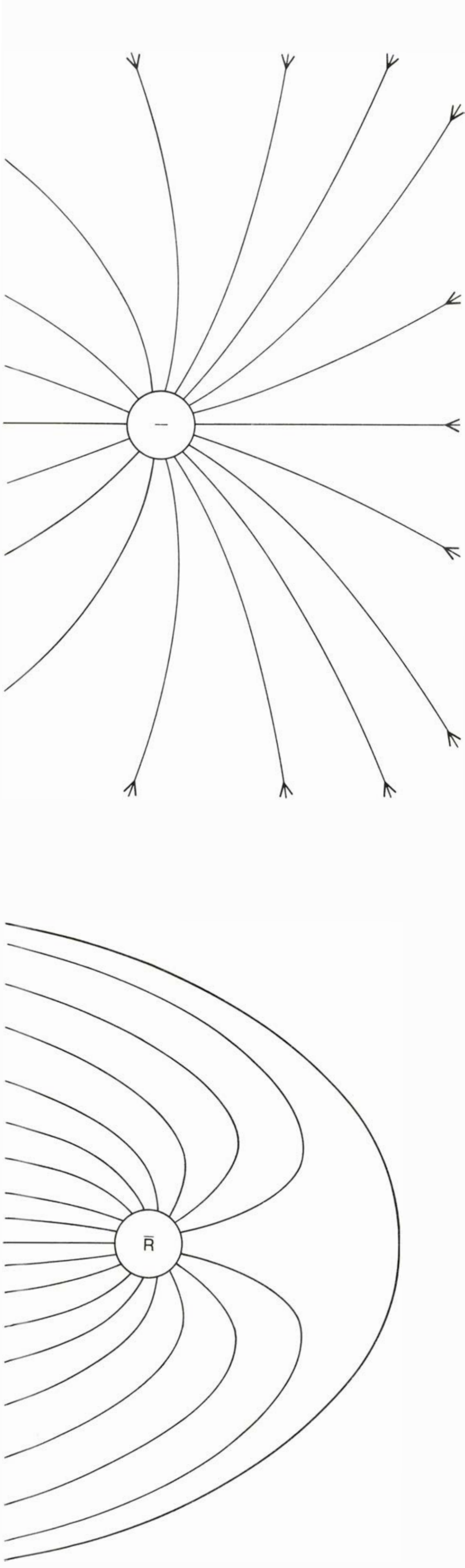
A finales de la década de los 60 la consideración de los quarks como componentes fundamentales de todos los hadrones fue confirmada de una forma rotunda por una serie de experimentos realizados por grupos dirigidos por Jerome I. Friedman y Henry W. Kendall, del MIT, y por Richard E. Taylor, del Stanford Linear Accelerator Center (SLAC). Se dispersaron inelásticamente electrones de alta energía por protones y el resultado de la dispersión indicó la existencia de partículas dentro de los protones. Estas partículas se comportaban tal como se suponía que lo hacían los quarks, y así se identificaron como quarks. El experimento de dispersión no sólo mostraba la realidad física de éstos sino también su naturaleza puntual. Por tanto, todos los componentes fundamentales de la materia, tanto los quarks como los leptones, parecen actuar como puntos sin dimensión.

El hecho de que tres quarks forman

LINEAS DE FUERZA ELECTRICAS (*arriba*) generadas por cargas eléctricas opuestas. Son curvas en el espacio que indican la dirección de la fuerza eléctrica a la que una carga positiva estaría sujeta si se le colocara sobre las curvas. El número de líneas en una región espacial dada indica la magnitud de la fuerza; cuanto mayor es el número de líneas, mayor es la fuerza. Las líneas de campo eléctrico llenan todo el espacio, pero lejos de las cargas la densidad de líneas de campo es baja y, por tanto, la fuerza también lo es. Líneas de fuerza de "color" (*abajo*) unen el par quark-antiquark de un hadrón genérico: cualquier partícula subatómica sujeta a la interacción llamada fuerza fuerte y que muestre signos de una estructura interna. Se designa como color una propiedad de los quarks que no tienen nada que ver con el color visual. Cada clase de quark aparece en tres colores: "rojo", "azul" y "amarillo". Los antiquarks tienen anticolores: "antirrojo", "antiazul" y "antiamarillo". Todos los hadrones son "blancos", es decir, son promedios incoloros de los tres colores y los tres anticolores. Los hadrones de dos quarks constan de un quark de un color, digamos rojo (*R*), y un antiquark del anticolor correspondiente, digamos antirrojo ($\bar{R}$). En el modelo en bolsa de los hadrones, los quarks están confinados en una bolsa o burbuja dentro del hadrón, y así las líneas de campo de color también están confinadas. El hecho de que las líneas sean paralelas intensifica la fuerza que actúa entre las cargas debido a que todas las líneas empujan en la misma dirección. Cuando en 1963 se introdujo el modelo de los quarks no se daba por supuesto su existencia real, sino que se pensó serían constructos mentales que ordenarían los hadrones.


+
−
R
R̄


+
R



un protón da cuenta de su carga fraccionaria (el protón, compuesto de dos quarks "arriba", "up", cada uno de carga 2/3 y un quark "abajo", "down", de carga −1/3). Además, la introducción de un nuevo número cuántico, o propiedad física, llamado "color" hace posible reconciliar el comportamiento de los quarks con el principio de exclusión. Sin embargo, como no se ha visto todavía ningún quark en estado libre, los quarks son aún objetos peculiares. Es relativamente fácil arrancar un átomo de una molécula, un electrón de un átomo o un protón de un núcleo atómico. Sólo se necesita una cantidad de energía relativamente pequeña para liberar cualquiera de estas partículas constituyentes. Enormes cantidades de energía no han permitido arrancar un quark siquiera de un hadrón.

Al principio, algunos físicos creyeron que se podría liberar un quark tan pronto como los aceleradores de partículas alcanzaran energías más altas. Esto no ha sucedido y en la actualidad se piensa que los quarks están permanentemente confinados en el interior de los hadrones. En la última década, los teóricos han propuesto varios modelos de confinamiento de quarks. El modelo de la bolsa es sólo uno de ellos, pero está ganando adeptos porque puede estar en estrecha relación con lo que la mayoría de teóricos consideran ser el mejor candidato a la teoría de las interacciones fuertes, la teoría de la cromodinámica cuántica, o dinámica de color.

Quiero introducir el modelo de la bolsa invocando una analogía sugerida por uno de mis estudiantes graduados, David Shalloway. Imaginemos un mundo cuya atmósfera consista en agua, no en aire. Podría gastarse energía haciendo hervir el agua, creando burbujas de vapor de agua. Dentro de las burbujas, las moléculas de vapor se comportarían, en primera instancia, como las partículas de un gas ideal que interaccionaran débilmente. En ese mundo hipotético, el vapor de agua podría existir sólo dentro de las burbujas. Una molécula de vapor de agua que, desde una burbuja, volviese al líquido no sería ya vapor de agua; cesaría de presentar las propiedades de una partícula de gas ideal.

El concepto de materia hadrónica en el modelo de la bolsa es parecido. Los hadrones son burbujas en el espacio que contienen quarks. En el interior de la burbuja los quarks se mueven libre e independientemente, igual que las moléculas de vapor de agua se mueven dentro de una burbuja. Si se intenta arrancar un quark de la burbuja aparece entre los quarks la interacción fuerte confinante. Igual que las moléculas de vapor de agua, los quarks sólo pueden existir en las burbujas que los contienen. En otras palabras, los hadrones se comportan como burbujas de vapor de quark ideal inmersas en una atmósfera de espacio ordinario en donde los quarks no pueden existir.

De la misma manera que hervir agua precisa cierta cantidad de calor latente, hace falta energía para "hervir" un vacío, es decir, para convertir energía en materia. De hecho, los aceleradores de partículas pueden considerarse hervidores de vacío. Se deposita una fracción grande de la energía cinética de las partículas aceleradas en un pequeño blanco. La mayor parte de la energía se emplea para formar quarks y antiquarks, que se combinan para constituir los hadrones hallados en los productos de la reacción. En el modelo de bolsa, algo de la energía también va al vacío, desplazando el espacio ordinario a fin de hinchar las burbujas que contienen los quarks. La cantidad de energía necesaria para hinchar una burbuja se denomina constante de la bolsa (B), y tiene un valor de alrededor de 55 millones de electronvolt por fermi cúbico (un fermi equivale a 10^{-15} metros).

Si una partícula está confinada en una región del espacio que mide Δx en cualquier dirección, según el principio de incertidumbre de la física cuántica el momento de la partícula tiene una indeterminación de $\hbar/\Delta x$, donde $\hbar$ es la constante de Planck dividida por 2π. Esto significa que, cuando los quarks están confinados en una bolsa de tamaño finito, se mueven con un momento del orden de $\hbar/\Delta x$, y, por tanto, ejercen una presión en la superficie interior de la bolsa. Para un cierto valor de equilibrio, la presión de los quarks en la superficie interior compensa la presión de confinamiento B que mantiene la materia dentro de la bolsa. Este valor de equilibrio

SEPARACION DE LAS CARGAS que afecta a la fuerza entre las mismas. Cuando dos cargas eléctricas opuestas (*arriba*) están separadas, las líneas de campo se curvan más. La fuerza eléctrica se hace más débil porque sólo la componente de las líneas de campo según la dirección que une las cargas contribuye a la fuerza resultante. La fuerza disminuye como el cuadrado de la separación entre las cargas (Ley de Coulomb). Cuando dos quarks (cargas de color) están separados (*abajo*), las líneas no se curvan sino que continúan empujando en la misma dirección. Por ser paralelas las líneas de campo de color, la fuerza de color entre un quark y un antiquark es una constante.

determina el tamaño de los hadrones. La presión del espacio ordinario que compensa la "presión del principio de incertidumbre" del vapor de quark recuerda bastante a la presión del agua que compensa la presión del vapor de agua en la burbuja.

Puede parecer que el problema del confinamiento de los quarks se ha resuelto de manera trivial postulando simplemente que los quarks están atrapados en bolsas o burbujas. Una bolsa que encerrara sólo un quark sería de hecho un quark, si bien "gordo". Tal partícula tendría la misma carga y demás propiedades del quark; así pues, el modelo de la bolsa, tal como lo he venido presentando hasta ahora, no aduce ninguna razón de por qué no se ha observado un quark aislado. Para explicar el confinamiento de los quarks es necesario introducir la propiedad dinámica del color. Cada sabor de quark se presenta en tres colores: "rojo", "azul" y "amarillo". Los antiquarks tienen anticolores: "antirrojo", "antiazul" y "antiamarillo". El color del quark no tiene nada que ver en absoluto con el color visual. Se usa la palabra color porque la manera como quarks de diferentes colores se combinan en mecánica cuántica recuerda la manera en que se combinan los colores de la visión. El mejor modo de imaginarse el color (y también el sabor) consiste en imaginarlo simplemente como una variable adicional asociada a los quarks. Del mismo modo que un quark tiene un momento angular de espín de $1/2\ \hbar$ que puede orientarse hacia arriba o hacia abajo según cualquier eje espacial, los quarks tienen sabores y colores que también pueden tener valores discretos.

Todos los hadrones son "blancos", es decir, son promedios incoloros de los tres colores. En términos aproximados, un barión ordinario consiste en un quark rojo, otro azul y un tercer amarillo, ya que la mezcla de estos colores es el blanco; los mesones ordinarios constan de un quark de un color y un antiquark del anticolor correspondiente. Una manera de establecer la verdad acerca del confinamiento de los quarks consiste en mostrar que la propiedad del color nunca puede observarse, porque observar un color sería ver un quark libre. Hoy se cree que la interacción fuerte se basa en el color de los quarks, del mismo modo que la interacción electromagnética se basa en la carga eléctrica. Las otras interacciones (la electromagnética, la débil y la gravitacional) son todas "ciegas al color".

Los efectos dinámicos del color de los quarks pueden aclararse mejor a través de una analogía con un concepto mucho más antiguo: la carga eléctrica. Michael Faraday introdujo la imagen de una carga como una fuente de líneas de fuerza eléctricas. Una línea de fuerza es una curva en el espacio que indica la dirección de la fuerza eléctrica a que estaría sujeta una carga positiva si se la colocara sobre la curva. El número de líneas en una región espacial dada indica la magnitud de la fuerza; a mayor número de líneas, más intensa es la fuerza. En la imagen de Faraday, las líneas de fuerza empiezan en las cargas positivas y acaban en las cargas negativas, porque una carga positiva es repelida por su propia clase de carga y atraída por su opuesta. El número de líneas que empiezan o acaban en una carga es una medida de la magnitud de la carga. Por tanto, cargas grandes generan fuerzas eléctricas fuertes.

El color de los quarks puede imaginarse como una clase de carga. En electrodinámica hay sólo un tipo de carga (que puede ser positiva o negativa),

		LEPTONES		QUARKS	
PRIMERA FAMILIA	PARTICULA	*e* (ELECTRON)	ν_e (NEUTRINO ELECTRON)	*u* (ARRIBA)	*d* (ABAJO)
	CARGA	−1	0	2/3	−1/3
	MASA	$.5 \times 10^{-3}$	0? ($<.6 \times 10^{-7}$)	$\sim .5 \times 10^{-2}$	$\sim 1 \times 10^{-2}$
	POSTULADO	——	1929	1963	1963
	DESCUBIERTO	1898	1954	——	——
SEGUNDA FAMILIA	PARTICULA	μ (MU)	ν_μ (NEUTRINO MUONICO)	*c* (ENCANTADO)	*s* (EXTRAÑO)
	CARGA	−1	0	2/3	−1/3
	MASA	.11	0? ($<.57 \times 10^{-3}$)	~ 1.7	~ .3
	POSTULADO	——	1957	1970	1963
	DESCUBIERTO	1936	1962	1974	1947
TERCERA FAMILIA	PARTICULA	τ (TAU)	ν_τ (NEUTRINO TAUONICO)	*t* (CIMA O VERDAD)	*b* (FONDO O BELLEZA)
	CARGA	−1	0	2/3	−1/3
	MASA	1.78	0? (<.25)	?	~ 5
	POSTULADO	——	1975	1977	——
	DESCUBIERTO	1975	?	19??	1977
¿CUARTA FAMILIA?		?	?	?	?
¿QUINTA FAMILIA?		?	?	?	?
. . . ?		?	?	?	?

FAMILIAS DE PARTICULAS ELEMENTALES. Cada una de ellas consta de dos clases de leptones (partículas que no están sujetas a la fuerza fuerte y que no muestran signos de estructura interna) y dos "sabores", o clases, de quarks, cada uno en tres colores distintos. Igual que los leptones, los quarks parecen actuar como puntos sin dimensiones. La carga se mide en unidades de la carga del protón, y la masa se da en términos de su energía equivalente en miles de millones de electronvolt (GeV). Aunque no se ha visto ningún quark aislado, se ha anotado una fecha de descubrimiento para cada sabor, que indica cuándo se descubrió una partícula que estaba formada exclusivamente por aquel sabor. El esquema de clasificación en familias se basa en el hecho de que los quarks de una familia dada interaccionan entre sí de la misma manera que lo hacen los quarks de cualquier otra familia. Los leptones se han asignado a las familias en orden de masa creciente. Se conocen tres familias y pueden descubrirse otras.

pero en cromodinámica hay varios tipos, que corresponden aproximadamente a las combinaciones de color permitidas. La fuerza de color actúa no sólo moviendo los quarks, sino también cambiando su color, de manera que el propio campo de color debe llevar color. (Un campo eléctrico actúa sólo moviendo las cargas eléctricas, pero no cambia ni su signo ni su magnitud; a este fenómeno se le dio generalidad matemática en una teoría de electrodinámica desarrollada por C. N. Yang, de la Universidad estatal de Nueva York en Stony Brook, y Robert L. Mills, de la Universidad estatal de Ohio.) Que los quarks interaccionen de esta forma es coherente con los resultados de los experimentos en que electrones de alta energía se dispersan de un modo inelástico por protones, los mismos experimentos que revelaron que los protones están formados por quarks. Esta coherencia fue señalada por David J. Gross y Frank Wilczek, de la Universidad de Princeton, y H. David Politzer, del Instituto de Tecnología de California. Afortunadamente no es necesario entender el mecanismo del cambio de color para apreciar la relación entre el color del quark y el confinamiento del mismo.

	NOMBRE	SABORES DE LOS QUARKS	ESPINES	CARGA	MASA (GEV)
MESONES	π^+	$u\bar{d}$	↑ ↓	+1	.14
	K^+	$u\bar{s}$	↑ ↓	+1	.50
	ρ^+	$u\bar{d}$	↑ ↑	+1	.76
	ϕ	$s\bar{s}$	↑ ↑	0	1.04
	J/ψ	$c\bar{c}$	↑ ↑	0	3.10
	D^{0*}	$c\bar{u}$	↑ ↑	0	2.00
	S^*	$s\bar{s}$ O $(u\bar{u} + d\bar{d})$	↑ ↓ ↑ ↓	0	1.00
	γ	$b\bar{b}$	↑ ↑	0	9.46
BARIONES	p^+	uud	↑ ↑ ↓	+1	.938
	n	ddu	↑ ↑ ↓	0	.940
	Δ^{++}	uuu	↑ ↑ ↑	+2	1.2
	Σ^+	uus	↑ ↑ ↓	+1	1.4
	Ω^-	sss	↑ ↑ ↑	−1	1.65

CONTENIDO EN QUARKS de algunos hadrones importantes. Los quarks conocidos aparecen en cinco sabores llamados "up" o "arriba" (*u*), "down" o "abajo" (*d*), "strange" o "extraño" (*s*), "charmed" o "encantado" (*c*) y "bottom" o "fondo" (*b*). Los antiquarks se designan por $\bar{u}$, $\bar{d}$, $\bar{s}$, $\bar{c}$ y $\bar{b}$. Los hadrones se han dividido en dos clases en función de cómo se desintegran, a saber: bariones y mesones. Los bariones son hadrones que se desintegran finalmente en un protón; todos están formados por tres quarks. Los mesones son hadrones que se desintegran completamente en fotones y leptones o bien en pares protón-antiprotón; la mayoría de mesones están formados por un único par quark-antiquark. Las flechas en la columna de espín (*hacia el centro*) indican la dirección en la que los quarks están girando.

Permítaseme continuar la analogía entre la carga eléctrica y la carga de color [*véanse las ilustraciones de las páginas precedentes*]. Nos puede ayudar la imagen de una línea de fuerza que empieza en una carga y acaba en otra como una especie de banda de goma que conecta las cargas, porque la tensión a lo largo de la línea de fuerza es en parte responsable de la fuerza atractiva entre las cargas. Las líneas adyacentes ejercen una presión entre sí, de manera que las líneas no se mezclan para formar una única línea recta que conecte las cargas. La fuerza neta entre las cargas es la suma de las fuerzas asociadas con cada línea. Cuando las cargas eléctricas se separan las líneas se curvan más. La fuerza se hace más debil porque sólo influye la componente de las líneas del campo según la recta que une las cargas. La fuerza disminuye como el cuadrado de la separación entre las cargas. Esta relación inversa al cuadrado de la distancia se conoce como ley de Coulomb.

Una diferencia importante entre un campo eléctrico ordinario y el campo de color tal como se describe en el modelo de la bolsa estriba en que las líneas de fuerza de color que conectan los quarks no existen en todas partes, sino que están confinadas en la bolsa. En la superficie de la bolsa, donde no hay quarks (y por tanto no hay cargas de color), las líneas no pueden empezar o acabar; y como las líneas no pueden salir de la bolsa, deben ser tangentes a su superficie. En la superficie, la presión hacia fuera de las líneas de campo de color no queda compensada por líneas adyacentes, ya que no hay tales, sino por la presión de confinamiento de la bolsa *B*. En consecuencia, las líneas de campo quedan paralelas unas a otras.

El hecho de que las líneas sean paralelas intensifica la fuerza que actúa entre las cargas, porque todas las líneas empujan en la misma dirección. A medida que los quarks (cargas de color) se separan, las líneas no se curvan sino que continúan empujando en la misma dirección. El número de líneas permanece también invariante. Por ser las líneas paralelas se sigue, en consecuencia directa, que la fuerza de color entre un quark y un antiquark es una constante. La magnitud de la fuerza es independiente de la distancia entre los quarks, excepto cuando están extremadamente juntos.

La fuerza constante entre quarks de color opuesto es lo que los mantiene permanentemente unidos. La magnitud de la fuerza, que es la tensión resultante en las líneas de campo de color que conectan las cargas, alcanza unas 15 toneladas. (Como contraste, la fuerza eléctrica del protón sobre el electrón en un átomo de hidrógeno es de unas 10^{-11} toneladas.) La fuerza de confinamiento es verdaderamente fuerte. Los quarks están confinados porque la constancia de la fuerza entre un quark y un antiquark implica que se necesita una cantidad arbitrariamente grande de energía para separarlos. Por ejemplo, para separar el quark y el antiquark de un mesón una pulgada se requeriría la energía necesaria para formar 10^{13} pares protón-antiprotón.

Los únicos agregados de quarks que tienen una masa finita son aquellos en que todas las líneas de campo de color acaban dentro de un volumen finito. Como las líneas deben empezar y acabar en cargas de color, tales agregados deben ser combinaciones incoloras de quarks. En otras palabras, todos los hadrones son blancos, de acuerdo con lo que se observa. El problema del confinamiento de los quarks se ha resuelto reduciéndolo al confinamiento de color. El que los colores estén confinados es una consecuencia del modelo en bolsa. La manera como los colores se combinan explica el hecho de que el barión consista en sólo tres quarks y el mesón conste de sólo un único par quark-antiquark. El agregado blanco más sencillo se compone de un

quark y un antiquark de colores opuestos. El siguiente en simplicidad consiste en un quark rojo, otro azul y un tercer amarillo. Es imposible hacer una combinación incolora de quarks a no ser que el número de quarks menos el número de antiquarks sea cero o bien un múltiplo de tres.

La cromodinámica, en unión con el modelo de la bolsa, puede explicar muchos aspectos del comportamiento de los hadrones que antes parecían paradójicos. Por ejemplo, si un quark y un antiquark estuvieran girando, se alejarían uno de otro como dos patinadores que describieran un círculo sujetos a los extremos de una cuerda elástica, con la salvedad de que, en el caso de los quarks, la cuerda consiste en líneas de color que conectan las cargas de color opuestas. Cuando se hace girar el sistema, su momento angular aumenta, las líneas de color se alargan y se transmite energía al campo de color que las líneas representan. Como resultado, la masa de sistema aumenta. La relación que predice el modelo entre el momento angular de espín y la masa ha sido verificada experimentalmente para todos los hadrones [*véase la ilustración de la página anterior*]. El descubrimiento experimental de esta relación fue la motivación principal de otro modelo de confinamiento de los quarks, el modelo de la cuerda, en el que dos quarks no pueden separarse porque están atados por una "cuerda". El modelo de la bolsa proporciona una explicación física de la naturaleza de la cuerda.

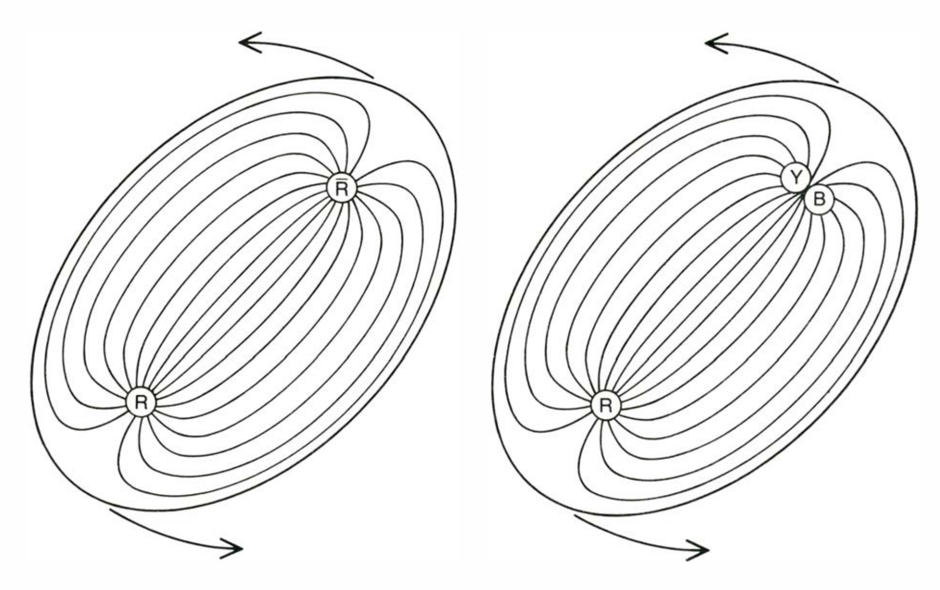
La deducción de la relación entre la masa y el momento angular obtenida a partir del modelo de la bolsa indica por qué la misma relación es aplicable a mesones y bariones. La relación depende sólo de la carga de color de los quarks en los extremos de la partícula. Como un barión es blanco, los quarks de uno de sus extremos tienen una combinación cromática que es opuesta al color del quark del otro extremo. En otras palabras, deben tener el mismo color resultante que el antiquark en el mesón.

El concepto de color puede también dar cuenta de la producción de hadrones. Un electrón y un positrón (un electrón cargado positivamente) pueden crearse como un par y, por tanto, también pueden hacerlo un quark y un antiquark. Tal proceso exige un "campo polarizante" que actúe para separar las cargas de color de un par de quark-antiquark en un mesón. Un nuevo par aparecerá espontáneamente en el campo de color del quark y el antiquark del mesón. Las líneas de color unen el quark del mesón con el antiquark recién creado y el antiquark del mesón con el recién formado quark. Ambos pares quark-antiquark son incoloros, de manera que las líneas de color no los conectan. Por no actuar una fuerza constante entre los pares, éstos pueden separarse. En otras palabras, la interacción fuerte puede transformar un mesón en otros dos, por ejemplo un mesón ϱ en dos piones.

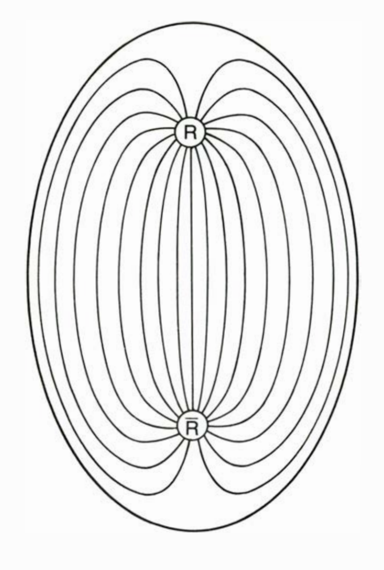
El concepto de líneas de campo de color también ha permitido describir la autoaniquilación de un electrón y un positrón a muy altas energías. Si las partículas poseen una energía suficientemente grande pueden generar un campo electromagnético lo bastante fuerte como para polarizar el vacío y crear un par quark-antiquark de cargas opuestas. Si los quarks del par son de poca masa, se separan tan deprisa como se juntaron el electrón y el positrón, pero no necesariamente a lo largo del mismo eje. Ni que decir tiene que el quark y el antiquark están conectados por líneas de campo de color. A medida que el quark y el antiquark se separan, rápidamente se crean pares de quarks coloreados por el fuerte campo de color, que forman mesones. La presión de las líneas de campo de color confina al par quark-antiquark, que se separa en una región tubular del espacio de un grosor de unos 10^{-15} metros. De acuerdo con el principio de incertidumbre, la componente del momento del par según la dirección transversal del tubo resulta ser el equivalente a sólo unos pocos centenares de millones de electronvolt. Como el quark y el antiquark se separan con gran celeridad, la componente del momento a lo largo del eje definido por su movimiento es mucho mayor. En consecuencia, los hadrones que emergen se mueven principalmente según esta dirección. Tales "jets" de hadrones acaban de observarse en el SLAC y en otros laboratorios.

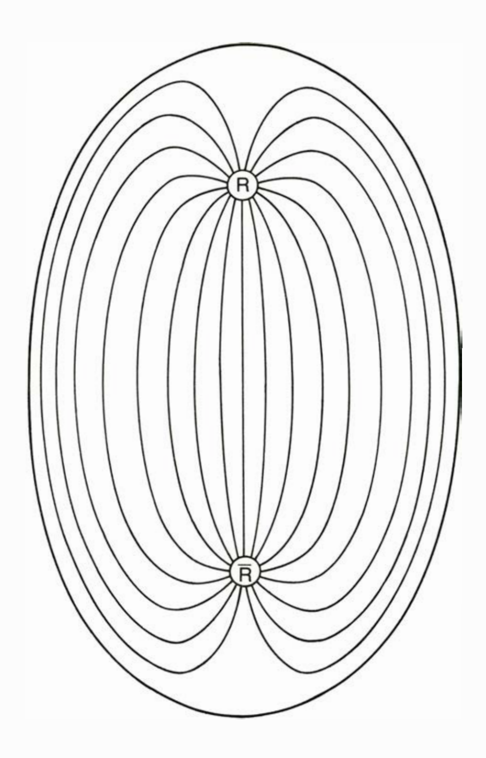
La electricidad y el magnetismo son fenómenos interrelacionados, ya que cuando las cargas eléctricas se mueven generan campos magnéticos. Como la carga de color se introdujo por analogía con la carga eléctrica, existe quizás una especie de magnetismo de color. De hecho lo hay, y una de sus consecuencias fue uno de los primeros éxitos de la cromodinámica. Esta consecuencia se descubrió independientemente por Howard Georgi, Alvaro de Rújula y Sheldon L. Glashow, de la Universidad de Harvard, y por un grupo de nosotros en el MIT. Es un hecho empírico bien conocido que las corrientes que fluyen en la misma dirección por dos alambres paralelos generan una fuerza magnética que tiende a unir los alambres. Igualmente, las corrientes que fluyen en direcciones opuestas generan una fuerza magnética que tiende a separar los alambres.

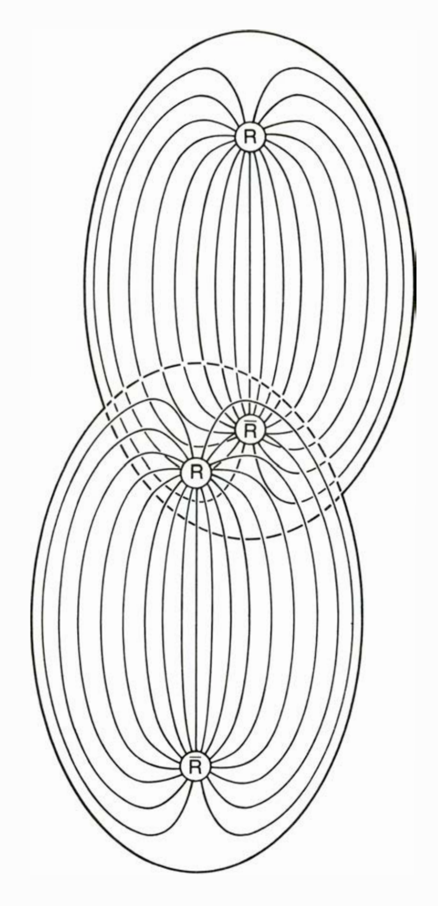
Como los quarks tienen momento an-



MESONES Y BARIONES tienen la misma relación entre momento angular y masa. El modelo en bolsa permite dar razón de ello. Si el quark y el antiquark de un mesón estuviesen girando (*izquierda*), se repelerían entre sí como dos patinadores que girasen en un círculo asidos al extremo de una cuerda elástica, con la salvedad de que en los quarks la cuerda consiste en las líneas de campo de color que unen cargas de color opuestas. Cuando se hace girar el sistema, el momento angular aumenta, las líneas de color se alargan y se traspasa energía al campo de color que las líneas representan. Como resultado, la masa del sistema aumenta. La relación depende sólo de la diferencia de color entre los quarks de cada extremo de la partícula. Como un barión (*derecha*) es blanco, los dos quarks (amarillo, *Y*, azul, *B*) de un extremo tienen un color resultante (antirrojo, *R̄*), *R* que es el opuesto del color (rojo, *R*) del quark del otro extremo. Así pues, los quarks tienen el mismo color que el antiquark del mesón.







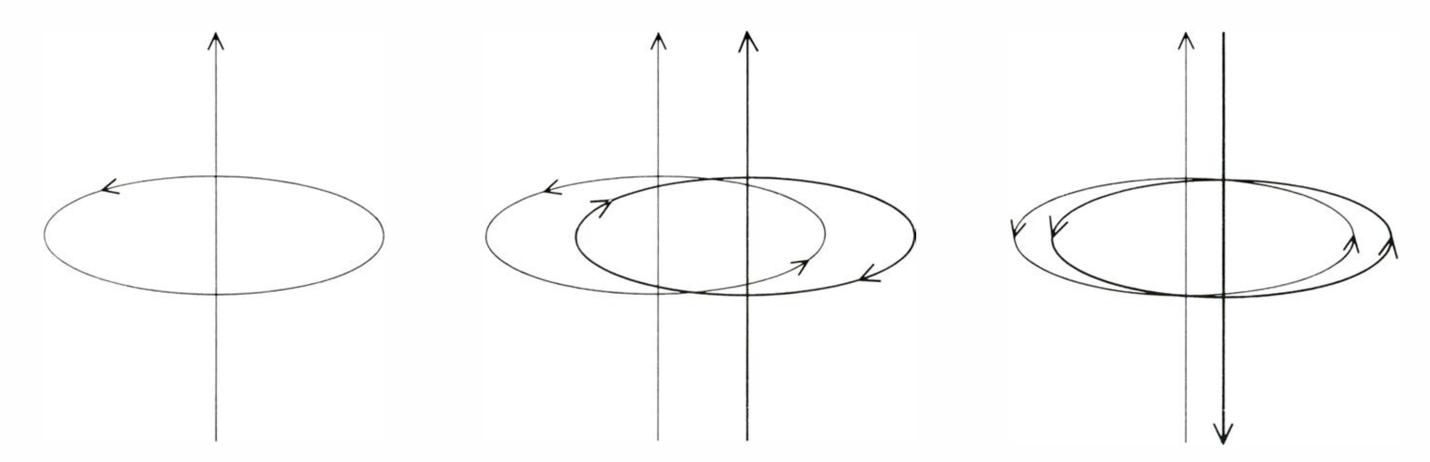
TRANSFORMACION DE UN MESON (*izquierda*) **en dos mesones: para ello se necesita un "campo polarizante" que actúe para separar las cargas de color del par de quarks del mesón** (*centro*). **Un nuevo par quark-antiquark aparece espontáneamente en el campo polarizante que hay entre el quark y el antiquark del mesón** (*derecha*). **Las líneas de color unen el quark del mesón con el antiquark recientemente creado y el antiquark del mesón con el quark recién creado. Ambos pares quark-antiquark son incoloros, de manera que las líneas de color no los unen. Debido a que no actúa una fuerza constante entre los pares (como lo hace entre el quark y el antiquark de cada par: rojo y antirrojo), pueden separarse y formar dos mesones.**

gular de espín, actúan como cargas de color en movimiento y, por tanto, generan un campo magnético de color. Un quark en giro puede imaginarse como una espira de corriente de color en la que una carga de color determinada circula en un plano que forma ángulo recto con la dirección del espín del quark. Como los quarks de un mesón son de color opuesto, si los espines son paralelos, la corriente de color fluye en direcciones opuestas, de modo que los quarks se repelen magnéticamente entre sí. Por otro lado, si los espines son antiparalelos, las corrientes de color fluyen en la misma dirección y la fuerza magnética de color atrae los quarks entre sí. Esto hace que las partículas que tienen los espines de los quarks alineados sean más pesadas que aquellas cuyos espines son opuestos. Esta consecuencia de la cromodinámica ha sido observada en todos los hadrones. Por ejemplo, el mesón ρ, que tiene los espines de los quarks paralelos, es más pesado que el mesón π, cuyos espines son antiparalelos.

El hecho de que el núcleo atómico esté formado por un conjunto de protones y neutrones, en vez de ser directamente un agregado de quarks, constituye a buen seguro una consecuencia de la magnitud de las fuerzas magnéticas de color. La energía de un conglomerado de quarks arriba ("up") y abajo ("down") es menor cuando los quarks se agrupan primero en paquetes incoloros (neutrones y protones), que a su vez se ligan entre sí (aunque más débilmente) para formar un núcleo atómico. La densidad de masa de los quarks en un hadrón es de sólo unas dos veces la densidad de masa de los nucleones (protones o neutrones) en el núcleo atómico. Esto indica que si de alguna manera la materia nuclear pudiera comprimirse de tal forma que la densidad de masa aumentase al menos en un factor dos, sería favorable energéticamente que todos los nucleones perdiesen su identidad individual pasando a formar un fluido de quarks. La transición de la materia nuclear a la materia de quarks no es, probablemente, una transición de fase (como sí lo es el cambio de un gas a un líquido cuando se baja la temperatura), sino que es una transición más gradual, tal como el cambio de un gas monoatómico a uno diatómico cuando se baja la temperatura.

Sería extremadamente difícil comprimir la materia nuclear en el laboratorio. Se cree, sin embargo, que la materia nuclear se comprime por las enormes fuerzas gravitacionales en una estrella de neutrones, el remanente frío de la explosión de una supernova. Si la densidad de la materia neutrónica comprimida supera la densidad de los quarks en el neutrón, los neutrones se fundirán para formar un fluido de quarks. Por tanto, la densa región central de una estrella de neutrones puede consistir de hecho en materia de quarks.

Esta posibilidad tiene consecuencias importantes para la estabilidad de la estrella de neutrones. La materia de la estrella está comprimida por la masa gravitacional, principalmente la masa de los neutrones. La resistencia a la compresión se debe a la interacción repulsiva de los neutrones y a la energía cinética de los neutrones fríos. Se cree que la energía cinética de los quarks atrapados da cuenta de la mayor parte de la masa de los neutrones; y así, la presión aumentaría enormemente si los neutrones se fundiesen y liberaran los quarks. Este efecto



UN CAMPO MAGNETICO DE COLOR se genera por un quark en movimiento (carga de color), del mismo modo que un campo magnético ordinario se genera por cargas eléctricas en movimiento. Un quark tiene momento angular de espín; por tanto, está siempre girando y puede ser tratado como una espira de corriente de color en la que circula una carga de color determinada (en sentido contrario a las agujas del reloj para espín hacia arriba) en un plano perpendicular a la dirección del espín del quark (*izquierda*). Como los quarks de un mesón son de color opuesto, si los espines son paralelos, las corrientes de color fluyen en direcciones opuestas (partículas formadas por quarks) de manera que los quarks se repelen magnéticamente entre sí (*centro*). Si los espines son antiparalelos, las corrientes de color fluyen en la misma dirección, de modo que la fuerza magnética de color atrae los quarks entre sí (*derecha*). Esto hace que los hadrones, cuyos espines estén alineados sean más pesados que los que tienen espines opuestos.

aumentaría de un modo notable la masa permitida a una estrella estable de neutrones. Se está trabajando en el análisis de la materia de quarks a fin de poder calcular esta masa permitida.

Como el modelo del campo de color describe tantas facetas de las propiedades e interacciones de los hadrones, los físicos de partículas esperan que el confinamiento de los quarks sea una consecuencia necesaria de la cromodinámica. ¿Es posible que el modelo en bolsa de los hadrones se deduzca a partir de la cromodinámica? La respuesta es afirmativa. Para entender cómo ocurre ello, consideremos una posible analogía con la superconductividad, a saber: la ausencia de resistencia eléctrica en un metal a una temperatura cercana al cero absoluto. Una fuerza magnética aplicada no puede penetrar un cuerpo superconductor. El campo magnético induce corrientes eléctricas en una fina capa de la superficie del cuerpo; estas corrientes dan lugar, a su vez, a un campo magnético que se opone al campo aplicado y es lo suficientemente grande como para cancelar el campo aplicado en el interior del cuerpo. Bajo ciertas condiciones, sin embargo, el campo magnético aplicado puede penetrar el cuerpo al crear en él una zona de conductividad normal, en la que las líneas de flujo magnético pueden quedar atrapadas. Este fenómeno se llama el efecto Meissner.

El efecto Meissner puede tener su análogo en cromodinámica. Quizás el vacío de la cromodinámica cuántica repele las líneas de campo de color igual que un superconductor repele las líneas de campo magnético. En este caso, los campos de color pueden existir sólo donde gastan energía para crear una zona de "vacío normal" (análoga a la zona de conductividad normal del efecto de Meissner) en el vacío circundante. La propiedad de este vacío envolvente de repeler el color haría el mismo papel que la presión de la bolsa en el modelo en bolsa, a base de mantener las líneas de campo de color confinadas en el interior de los hadrones. Charles G. Gross, Curtis G. Callan y R. Dashen, de Princeton, han hecho cálculos detallados que apoyan este modelo de vacío dual. Otro modelo de esta clase ha sido estudiado por Holger B. Nielsen y M. Minomiya, del Instituto Niels Bohr de Dinamarca. Estos modelos constituyen el frente del progresivo trabajo sobre el confinamiento de los quarks.

La cromodinámica cuántica y el modelo de la bolsa han tenido bastante éxito al proporcionar un mecanismo para el secuestro permanente de los quarks dentro de los hadrones y explicar, por tanto, muchos aspectos de la fuerza fuerte entre los quarks. El que este mecanismo tenga un éxito definitivo o no depende de cómo funcione cuando se le desarrolle para hacer cálculos más refinados y detallados de las propiedades de los hadrones. El fenómeno del confinamiento de los quarks es quizás el aspecto más paradójico del modelo de quarks. Una explicación definitiva del fenómeno eliminará la última sombra importante en los muchos éxitos de la hipótesis de los quarks.

Es ahora el momento de hacer el inventario de las partículas subatómicas que ha surgido del modelo de quarks. Se considera que los quarks son los constituyentes elementales de la materia, ya que parecen interaccionar como entes puntuales, pero ¿son realmente elementales? La historia de la ciencia proporciona muchos ejemplos de entes físicos que alguna vez se pensó que eran elementales pero que más tarde resultaron ser compuestos. Los átomos, los núcleos atómicos, los protones y los neutrones se creyeron indivisibles en algunas etapas de este siglo hasta que posteriores experimentos descubrieron sus partes constituyentes. ¿Es posible que los quarks sufran el mismo destino? ¿Y qué pasa con los leptones, que también se consideran elementales? ¿Habrá que buscarles también partes constituyentes?

La posibilidad de que los quarks o los leptones tengan una subestructura es descartada por la mecánica cuántica y la teoría de la relatividad. Se cree que los quarks y los leptones no superan los 10^{-17} metros. Para el electrón este límite superior de tamaño se deduce directamente de los experimentos. A partir del principio de incertidumbre, podemos calcular el momento de un ente constituyente hipotético confinado en una región del espacio de 10^{-17} metros de diámetro. La teoría de la relatividad indica que la energía cinética de los constituyentes sería mayor que su momento multiplicado por la velocidad de la luz. La energía cinética del constituyente confinado en una región de 10^{-17} metros de diámetro aumentaría la masa del quark o leptón del que forma parte; el constituyente menos energético contribuiría con una masa de 20 veces, al menos, la masa del protón. Por tanto, es extremadamente improbable que el electrón, por ejemplo, tenga constituyentes. Para que fuera así,

debería haber una cancelación accidental y casi perfecta de la energía de enlace de las partículas constituyentes con su energía cinética. En física estos accidentes suceden raras veces, de manera que es muy improbable que los electrones (igual que los otros leptones y quarks) sean compuestos en este sentido.

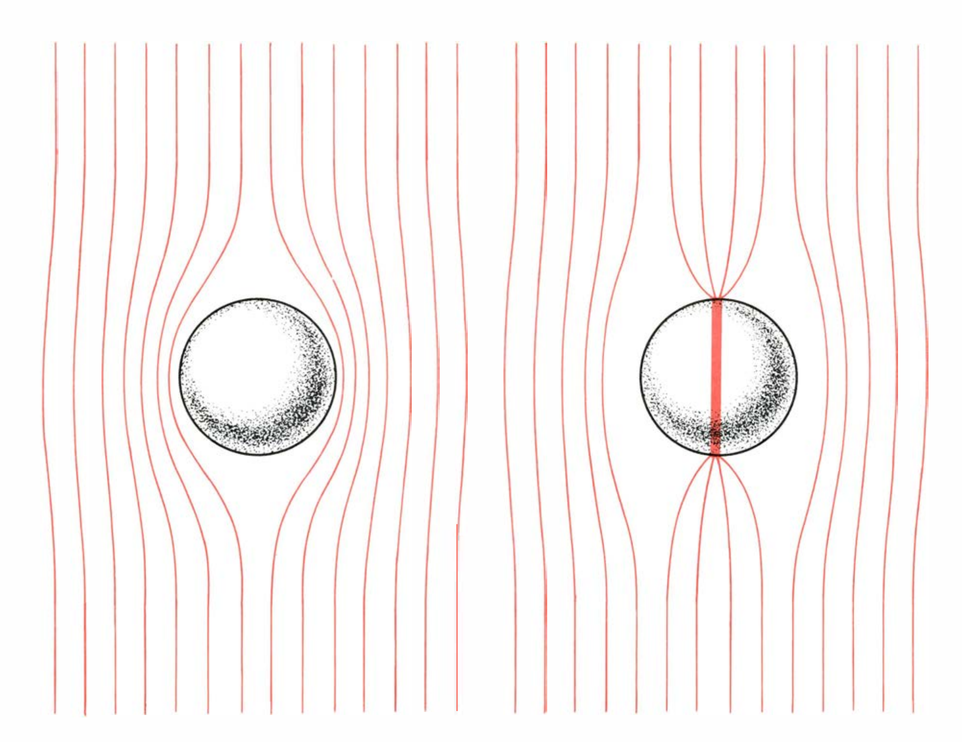
Si los quarks y leptones son realmente elementales, la naturaleza no puede continuar complicando la imagen de la materia con la creación de nuevas clases de partículas que sean incluso más fundamentales. En cambio parece estar complicando la imagen creando nuevos sabores de quarks y nuevas especies de leptones. Los quarks y leptones conocidos pueden clasificarse agrupándolos en familias. Cada familia consta de cuatro partículas cuya carga eléctrica total es cero, simplemente dos leptones (una partícula y su neutrino asociado) y dos sabores de quarks (en tres colores distintos: rojo, azul y amarillo). Hasta ahora se han identificado tres familias. La primera está formada por el electrón, el neutrino electrónico, el quark arriba ("up") y el quark abajo ("down"). La segunda está formada por el muon, el neutrino muónico, el quark "encantado" y el quark "extraño". La tercera está formada por la partícula tau, su neutrino, el quark fondo ("bottom") y el quark cima ("top"), no detectado todavía, si bien teóricamente esperado.

Los miembros de la primera familia desempeñan el papel más importante en los fenómenos que suceden naturalmente. Las otras partículas elementales, nacidas sobre todo en los aceleradores de partículas, tienen una existencia efímera y son responsables tan sólo de efectos extraordinariamente sutiles en la materia ordinaria. Se desintegran en miembros de la primera familia mediante la interacción débil.

Este esquema de clasificación se funda en que las cuatro partículas de una determinada familia interaccionan entre sí, de la misma manera que lo hacen las cuatro partículas de otra familia. Por ejemplo, dentro de una familia un sabor de quark puede transformarse en el otro por la interacción débil. [Esto sucede en la radiactividad ordinaria donde un quark arriba ("up") se cambia en un quark abajo ("down") o viceversa.] Un leptón puede transformarse asimismo en otro. Los quarks de distintas familias pueden interaccionar también, aunque esta interacción es considerablemente más débil. No se conocen casos de cambios de leptones de distintas familias, cambios de leptones en quarks ni de quarks en leptones.

	ESPINES PARALELOS			ESPINES ANTIPARALELOS		
MOMENTO ANGULAR ORBITAL	PARTICULA	CONTENIDO QUARKS	MASA (GEV)	PARTICULA	CONTENIDO QUARKS	MASA (GEV)
0	Δ^+	*uud*	1.23	*p*	*uud*	.94
0	Σ^{+*}	*uus*	1.38	Σ^+	*uus*	1.19
0	Ξ^{-*}	*dss*	1.54	Ξ^-	*dss*	1.32
0	ρ^+	$u\bar{d}$	.77	π^+	$u\bar{d}$	.14
1	$A_2{}^+$	$u\bar{d}$	1.31	B^+	$u\bar{d}$	1.23
2	g^+	$u\bar{d}$	1.69	A_3	$u\bar{d}$	1.64
0	K^{-*}	$u\bar{s}$	.89	K^-	$u\bar{s}$	.49
1	K^{-**}	$u\bar{s}$	1.42	"Q"	$u\bar{s}$	1.3
2	K^{-***}	$u\bar{s}$	1.78	L^-	$u\bar{s}$	1.77

EFECTOS DE MASA DEL MAGNETISMO DE COLOR para varios hadrones. La tabla indica claramente que, a igualdad de las otras características, una partícula en la que los espines de los quarks son paralelos tiene una masa mayor que otra cuyos espines sean antiparalelos. El momento angular orbital se da en unidades de la constante de Planck dividida por 2π. Cuando los quarks giran se separan uno de otro como resultado de la fuerza centrífuga a la que están sujetos en el sistema de referencia que gira; cuanto mayor sea el ritmo de rotación, mayor será la separación. La interacción magnética de color se debilita al aumentar la separación. Esto indica que a medida que el momento angular orbital aumenta (el ritmo de rotación crece) las masas de los estados paralelos y antiparalelos se igualan.



PROPIEDADES DE UN SUPERCONDUCTOR (un metal a temperaturas cercanas al cero absoluto que no tiene resistencia a la corriente eléctrica). Dichas propiedades pueden servirnos para establecer una analogía de cómo el confinamiento de los quarks puede deducirse de la cromodinámica, o dinámica del color. Un campo magnético aplicado puede penetrar un cuerpo superconductor (*izquierda*). El campo magnético induce corrientes eléctricas en una capa delgada de la superficie del cuerpo; las corrientes dan lugar, a su vez, a un campo magnético que se opone al campo aplicado, y es suficiente para cancelar el campo aplicado dentro del cuerpo. La única manera como el campo magnético aplicado no puede penetrar en el cuerpo es a base de crear una región de conductividad normal, en la que las líneas de flujo magnético pueden quedar atrapadas (*derecha*). El vacío en la cromodinámica cuántica puede expulsar las líneas de campo de color igual que un superconductor expulsa las líneas de campo magnético. En dicho caso los campos de color pueden sólo existir donde gastan energía para crear, en el vacío circundante, una región de "vacío normal" análoga a la región de conductividad normal. La propiedad de expulsar el color del vacío circundante haría la misma función de la bolsa en el modelo del mismo nombre, es decir, mantener las líneas de campo de color, y por tanto los quarks, confinados.

Todos los miembros de la primera familia habían sido identificados a finales de los años 60 cuando los quarks arriba ("up") y abajo ("down") fueron detectados dentro del protón y del neutrón. Las partículas de la segunda familia fueron descubiertas entre 1936 y 1974. El descubrimiento en 1974 del quark "encantado", el último miembro de la segunda familia, dio una tremenda fuerza al modelo de quarks. Cuatro años antes, la necesidad de la existencia del quark encantado había sido deducida a partir de consideraciones teóricas por Glashow, John Iliopoulos y Luciano Maiani. El quark fue posteriormente descubierto por Samuel C. C. Ting y sus colaboradores en el Brookhaven National Laboratory y por Burton D. Richter y sus colaboradores en el SLAC. La tercera familia apareció en escena en 1975 con el descubrimiento de la partícula tau. Aunque no se sabe con certeza si el neutrino tauónico es distinto del neutrino electrónico, la mayoría de los físicos de partículas creen que lo es. A menos que el quark cima ("top") sea extraordinariamente pesado, deberá descubrirse dentro de los próximos cinco años.

No hay ninguna razón teórica que impida el descubrimiento de aun más familias de partículas. Algunos teóricos creen que las teorías unificadas de las interacciones débiles y electromagnéticas pueden eventualmente dar indicios de esta posibilidad. A fines de los años 60, una de tales teorías fue propuesta independientemente por Steven Weinberg, de la Universidad de Harvard, y por Abdus Salam, del Centro Internacional de Física Teórica de Trieste. Su propuesta explica satisfactoriamente muchas propiedades de las interacciones débiles. En su forma actual, esta propuesta es totalmente compatible con el esquema de clasificación en familias, pero no ofrece ninguna explicación para la existencia de las familias.

Aunque otras propuestas de mayor alcance que incorporan la teoría de Weinberg y Salam proporcionan tal explicación, las mismas carecen de base empírica. Además, no se ha dado ninguna explicación clara de por qué las partículas de una familia dada tienen las masas que ostentan. Mientras que un modelo definitivo del confinamiento de los quarks y una confirmación de la naturaleza elemental de los quarks y los leptones cerraría un largo capítulo de la física de partículas, el de la búsqueda de los últimos constituyentes de la materia ordinaria, la posible proliferación de familias, está abriendo otro capítulo.

Los animales de Burgess Shale

Los fósiles de una formación rocosa del Canadá occidental constituyen una rica muestra de una comunidad animal del Cámbrico medio. Algunas especies son antepasadas de las que viven hoy en día, pero otras son únicas y extrañas

Simon Conway Morris y H. B. Whittington

Con mucho, los fósiles más numerosos que representan a los primeros seres vivos abundantes en la Tierra están constituidos por las partes duras de varios animales invertebrados marinos: conchas y caparazones y fragmentos similares de esqueleto externo. Esto supone un registro fósil desequilibrado. Por ejemplo, de los aproximadamente 30 tipos (phyla) de animales actuales, más de la mitad están constituidos por especies con pocas partes duras o sin ellas. En consecuencia, el origen filético de estos grupos carece en la mayoría de los casos de una documentación fósil.

Mas, por fortuna, la situación no es completamente desequilibrada. Se han descubierto unos pocos yacimientos geológicos que, como resultado de circunstancias excepcionales, contienen fósiles exquisitamente conservados de animales cuyo cuerpo era parcial o totalmente blando. En este artículo describiremos uno de tales yacimientos, el de Burgess Shale, del Canadá occidental. La edad elevada y la gran variedad de invertebrados marinos en Burgess Shale hace quizá de este yacimiento el mejor conocido. Además de describir la fauna de Burgess Shale intentaremos reconstruir el tipo de ambiente subacuático que estos organismos habitaban a principios del Paleozoico, hace unos 530 millones de años.

En el otoño de 1909 el entonces secretario de la Smithsonian Institution, Charles Doolittle Walcott, estaba buscando formaciones fosilíferas en la Columbia Británica. Siguiendo un sendero que recorría la pendiente occidental entre Wapta Mountain y Mount Field, en la parte meridional de la provincia, Walcott tropezó literalmente con un bloque de pizarra que había caído sobre el sendero desde la pendiente situada encima. Al examinar la roca fácilmente fisible, se sorprendió al encontrar impresiones fósiles de varios animales de cuerpo blando conservadas en sus capas. En una carta dirigida a un colega de Toronto, fechada el 27 de noviembre de 1909, indicaba escuetamente que había permanecido "unos cuantos días recolectando... en las inmediaciones de Field y encontré algunas cosas interesantes."

Walcott volvió al mismo lugar al año siguiente para prospectar pendiente arriba en busca del estrato de esquisto de donde había caído la roca que encontró. Su búsqueda tuvo éxito: halló dos afloramientos de esquistos fosilíferos separados por una distancia vertical de unos 20 metros. Excavó someramente en ambos; el afloramiento más bajo resultó ser el más rico de los dos. Envió al Distrito de Columbia miles de especímenes fósiles que obtuvo de lo que llamaba su Yacimiento de Filópodos. (Este término, poco utilizado por los paleontólogos contemporáneos, se refiere a ciertos artrópodos, o animales articulados, fósiles que probablemente son los antepasados de los crustáceos actuales.)

Como demostraron el propio trabajo de Walcott y décadas de estudio por parte de otros autores, los fósiles de Burgess Shale (literalmente, Esquisto de Burgess) incluyen una gran abundancia de invertebrados marinos: más de 120 especies. Algunas de ellas pertenecen al tipo Poríferos, las esponjas. Este tipo de animales primitivos es el único del subreino Parazoos, categoría superior a la del subreino Protozoos (unicelulares) pero inferior a la del subreino Metazoos (pluricelulares). Quizás otras 10 especies representan tipos de metazoos que eran desconocidos antes de ser descubiertos en Burgess Shale; no se les encuentra en ningún otro lugar del registro fósil. Las muchas otras especies que carecen de partes duras pueden adscribirse a un tipo u otro de metazoos con parientes actuales, de la siguiente manera:

Celentéreos: tipo (phylum) que incluye animales marinos vivientes como las medusas, las gorgonias y los corales. Las especies de celentéreos de Burgess Shale pueden ser hasta cuatro.

Equinodermos: tipo que contiene, entre otros, las estrellas de mar, los erizos de mar, las holoturias y los crinoideos. Se han reconocido por lo menos cuatro especies de equinodermos en Burgess Shale.

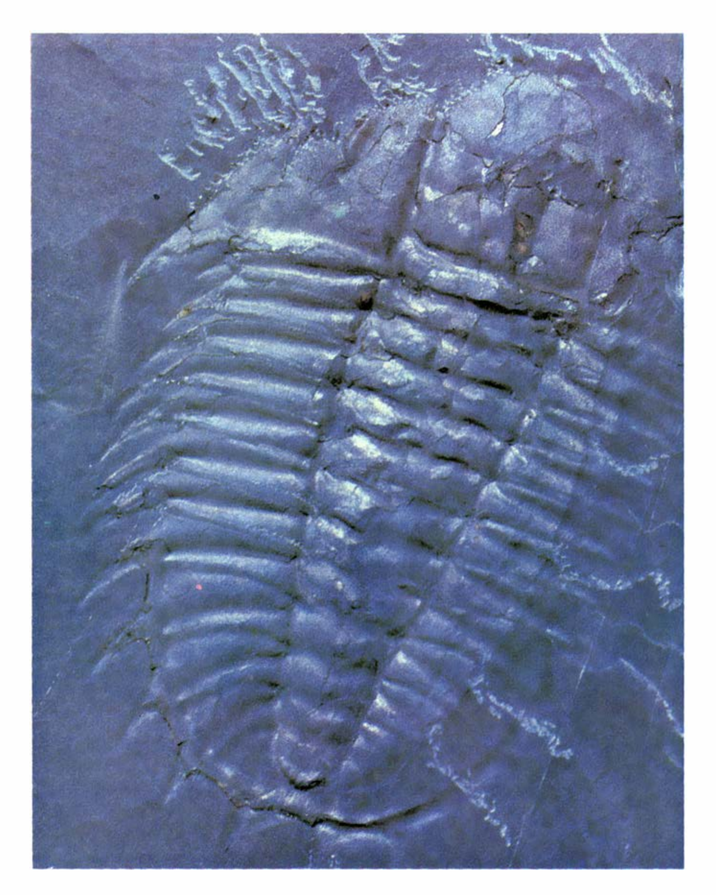
Moluscos: tipo que contiene, entre otros, las ostras y las almejas, los caracoles, los calamares y pulpos y los primitivos quitones (de la clase Anfineuros). Se han reconocido tres especies de moluscos en Burgess Shale.

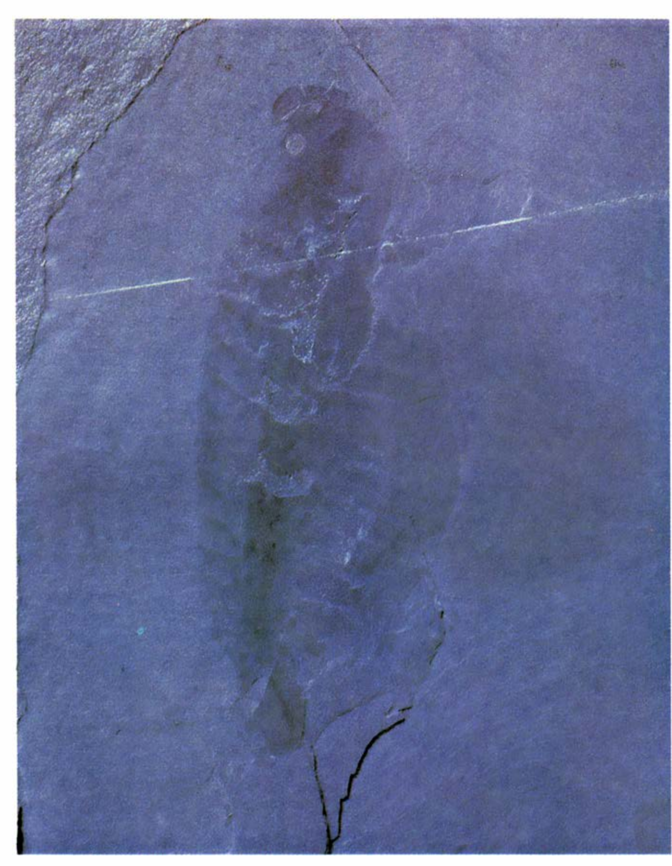
Artrópodos: tipo que incluye, entre muchísimos otros, langostas de mar, quisquillas, cangrejos y bellotas de mar (todos de la clase Crustáceos) y el animal terrestre, menos familiar, *Peripatus*, miembro de la case Onicóforos. Los artrópodos de Burgess Shale comprenden varios representantes de la clase Trilobites, extinguida desde muy antiguo; un animal semejante a *Peripatus* y que era acuático y no terrestre, y alrededor de otras 30 especies de artrópodos.

Priapuloideos: tipo menor de gusanos marinos no segmentados. El género actual *Priapulus* da su nombre al grupo. En los fangos de Burgess Shale prosperaron siete especies de estos excavadores bentónicos hoy poco conocidos.

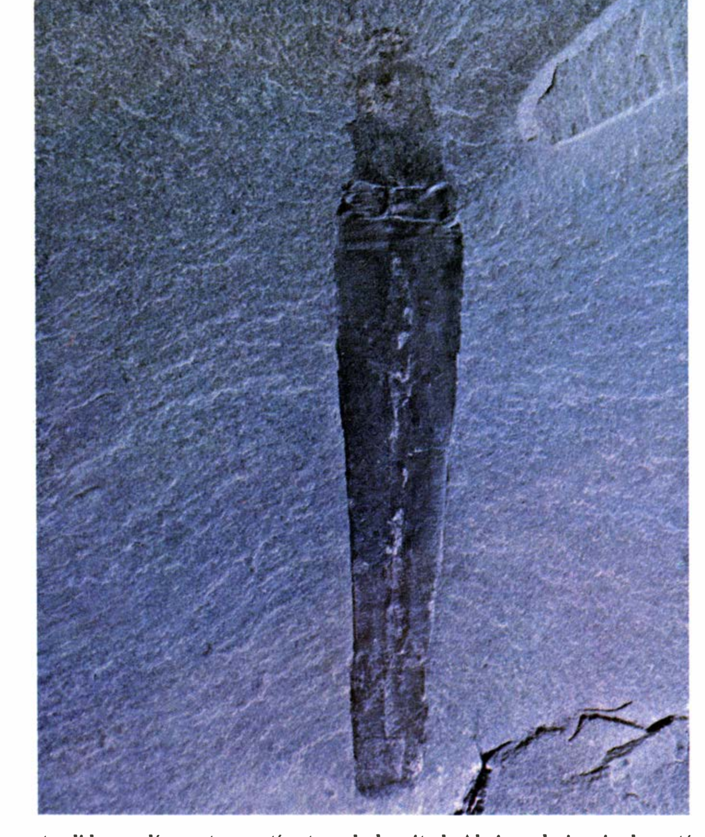
Anélidos: tipo que comprende las lombrices de tierra, las sanguijuelas y una clase menos familiar pero muy extensa de gusanos marinos, los Poliquetos. El tipo Anélidos está representado en Burgess Shale por seis especies.

Finalmente, en la fauna de Burgess Shale se encuentra también uno de los representantes invertebrados más primitivos de nuestro propio rincón del reino animal: el tipo Cordados. Entre sus representantes actuales (y además de los vertebrados) se cuentan las ascidias y *Amphioxus* o *Branchiostoma*, un pecu-



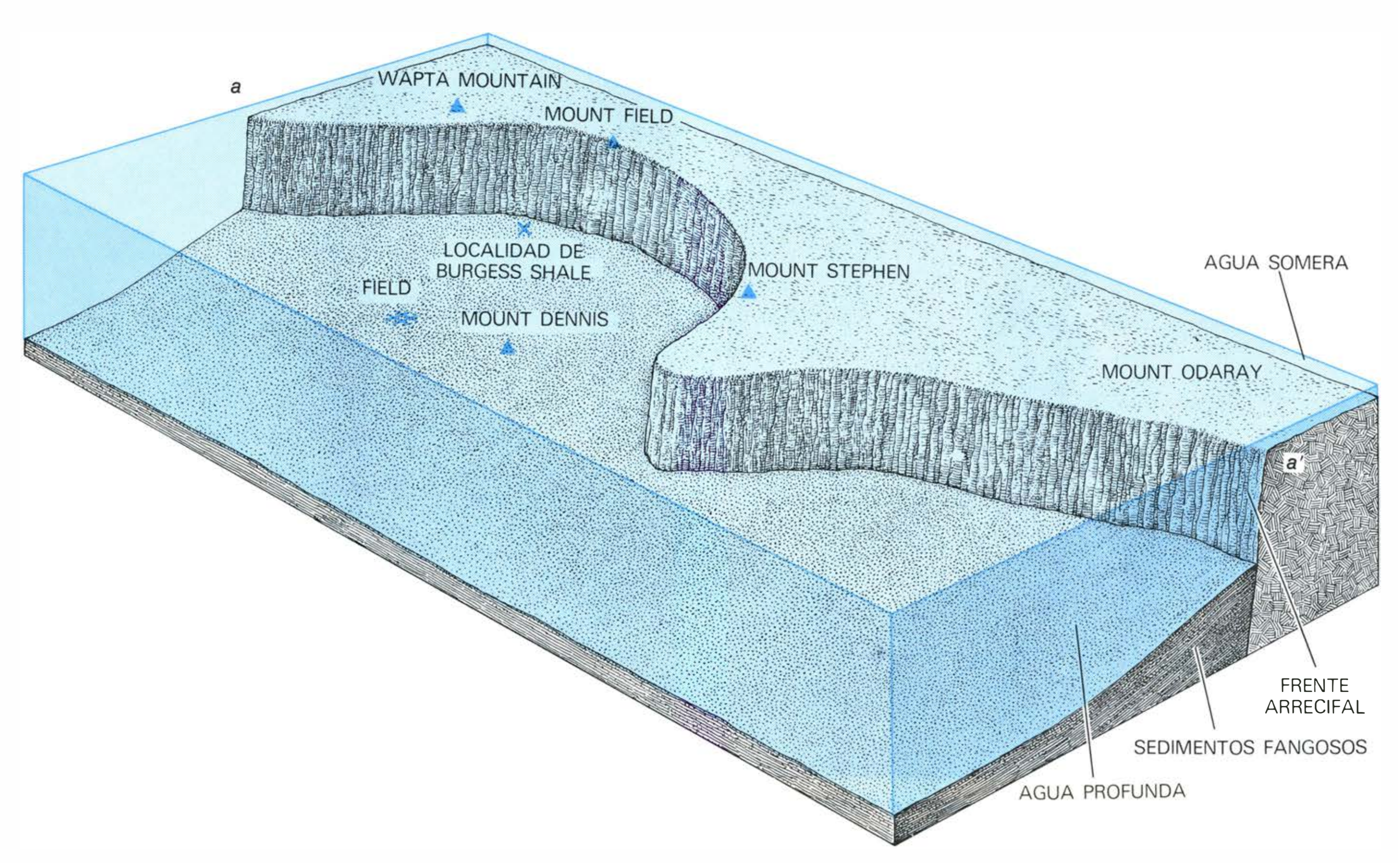






CUATRO ANIMALES que vivieron en el océano en el Cámbrico medio, hace unos 530 millones de años, se ven en estos fósiles. Arriba, a la izquierda, hay un trilobites, *Olenoides*, uno de los muchos animales cuya anatomía se ha conservado con un notable detalle en los limos que se solidificaron para formar Burgess Shale, en Columbia Británica. El espécimen tiene 5,5 centímetros de longitud. A diferencia de la mayoría de artrópodos, o animales de patas articulados, *Olenoides* tenía las extremidades no especializadas. Arriba, a la derecha, hay otros artrópodo de Burgess Shale, *Waptia*. Cuando este animal, que se alimentaba en el fondo, se hallaba extendido, medía cuatro centímetros de longitud. Abajo, a la izquierda, está *Opabinia*, una de las aproximadamente 10 especies animales encontradas en los esquistos que pertenecen a tipos filéticos anteriormente desconocidos. Tenía cinco ojos y dirigía su cuerpo de siete centímetros con una aleta caudal vertical mientras nadaba cerca del fondo marino en busca de comida. Abajo, a la derecha, se muestra uno de los muchos gusanos marinos no segmentados que vivían en el fondo marino; se trata de *Selkirkia*, del tipo Priapuloideos. Con su trompa evaginada medía cinco centímetros. Un grupo próspero en el Cámbrico, los Priapuloideos son raros en la actualidad.

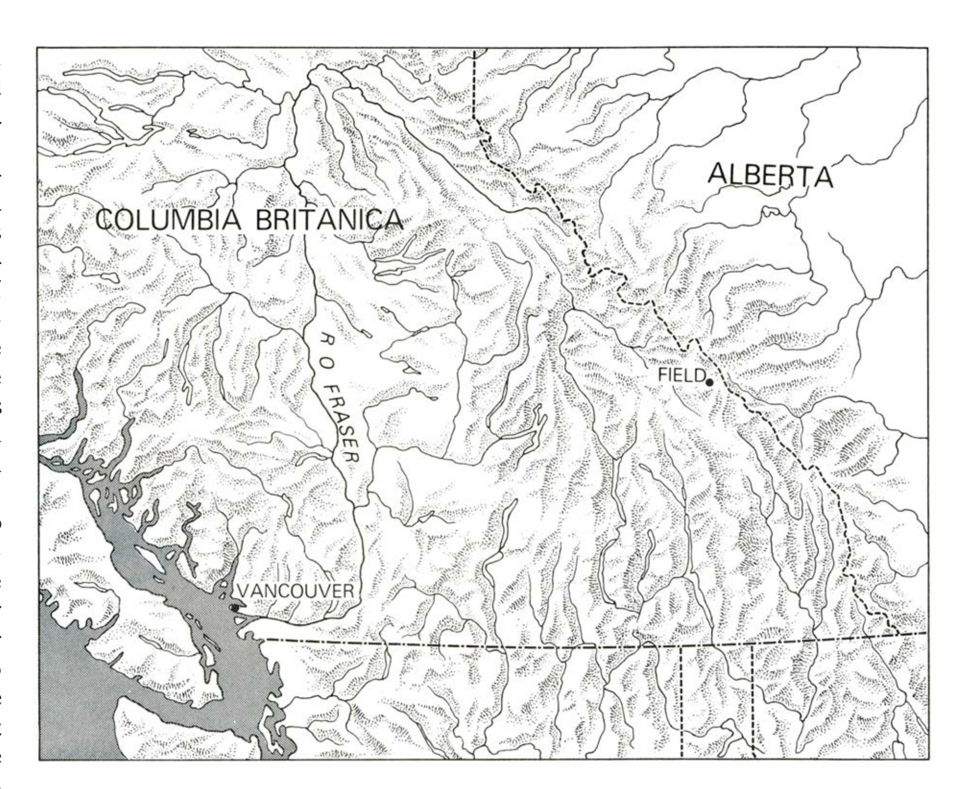
CANTERA INFERIOR, llamada Yacimiento de Filópodos por Charles D. Walcott, el primero en prospectar Burgess Shale; presenta en esta foto que mira al sur manchas de nieve invernal. Este afloramiento de esquistos y otro superior fueron prospectados de nuevo en busca de fósiles en 1966.



AFLORAMIENTO DE BURGESS SHALE, marcado con una *X* de color en este esquema de un paisaje marino del Cámbrico medio. Constituye una pequeña parte de un extenso yacimiento bentónico de limos acumulados al pie de un arrecife algal muy resguardado en el fondo de una bahía y que se elevaba unos 160 metros por encima de los limos más someros. El arrecife no sobrepasaba el nivel del mar, sino que estaba cubierto por aguas someras. La escala vertical en el esquema se ha exagerado por un factor de cinco, y la distancia (*a-a'*) de norte a sur a lo largo del arrecife es de unas nueve millas. Los levantamientos y divisiones posteriores dieron origen a los picos de las Rocosas a lo largo de la frontera de Alberta y de la Columbia Británica, indicados mediante triángulos de color. La reconstrucción del fondo marino y del arrecife se basa en el trabajo de I. A. McIlreath.

liar animal marino. Los cordados están representados en Burgess Shale por el género *Pikaia* y la especie única *P. gracilens.*

Esta fauna de cuerpo blando, notablemente bien conservada y que representa ocho tipos filéticos conocidos y 10 o más anteriormente desconocidos que prosperaron en el Cámbrico medio, tiene por sí misma un gran interés para los estudiosos del registro fósil. Pero además de este interés intrínseco, los invertebrados de Burgess Shale, con sus adaptaciones especializadas, poseen una importancia todavía mayor en la clarificación de la evolución temprana del reino animal. Los únicos animales de cuerpo blando más primitivos conocidos son representantes del Precámbrico tardío, hace de 700 a 600 millones de años, y son por tanto 70 millones de años anteriores, por lo menos, a los tiempos del Cámbrico Medio. Se trata de los animales de Ediacara, descubiertos por primera vez hace 30 años en las Ediacara Hills, de Australia meridional, y que desde entonces se han identificado en varios otros lugares del globo. La fauna de Ediacara contrasta fuertemente con la de Burgess Shale tanto por el tipo de animales representados (principalmente celentéreos) como por la limitada diversidad de especialización de estos animales primitivos.



FORMACION DE BURGESS SHALE, situada a unos 563 kilómetros al nordeste de Vancouver, cerca del pueblo de Field. En 1909, Walcott encontró accidentalmente esta formación rica en fósiles.

El acontecimiento que separa la fauna emprobecida de Ediacara de la fauna de Burgess Shale es una diversificación evolutiva explosiva de animales pluricelulares que tuvo lugar alrededor del comienzo de los tiempos cámbricos. Los fósiles de Burgess Shale nos permiten pues dirigir una ojeada única a los resultados de esta súbita adaptación de los Metazoos, y relativamente poco después de que tuviera lugar.

A pesar del trabajo realizado por Walcott y otros, persisten todavía lagunas significativas en lo que se sabe del paleoambiente de Burgess Shale y de cómo se preservó su fauna. Una apreciación más completa de estas lagunas estimuló una reinvestigación del lugar por parte del Geological Survey del Canadá, que empezó hace más de una década. Las autoridades del Yoho National Park en Columbia Británica y la administración de Parks Canada en Ottawa concedieron permisos especiales para recolectar material en los afloramientos de esquistos. Las canteras de Walcott se volvieron a abrir en 1966 y 1967, bajo la dirección de J. D. Aitken, del Geological Survey del Canadá. Posteriormente, nos llegó para su análisis en la Universidad de Cambridge tanto el nuevo material recolectado durante estas dos temporadas como una parte de la gran colección de Burgess Shale que Walcott había reunido hace 60 años.

¿En qué tipo de ambiente vivía la fauna de Burgess Shale? Los estudios llevados a cabo por I. A. McIlreath, de Petro-Canada, y por W. H. Fritz, del Geological Survey del Canadá, indican que los animales vivieron sobre o en el interior de un fondo fangoso en el que los sedimentos se habían acumulado en la base de un arrecife gigantesco. Esta estructura, constituida por material secretado por algas, se elevaba verticalmente varias decenas de metros a partir de una cuenca de aguas profundas que gradualmente se iba colmatando de sedimentos. Todavía pueden hallarse a lo largo y a lo ancho de la Columbia Británica afloramientos dispersos del frente arrecifal. Aparentemente, las aguas del fondo de la cuenca tenían una circulación limitada, eran ricas en sulfuro de hidrógeno y pobres en oxígeno. Los distintos invertebrados medraban allí donde los sedimentos fangosos se situaban a una altura suficientemente elevada contra el arrecife para librarse de las aguas estancadas del fondo, a unos 160 metros por debajo del nivel del mar.

Los sedimentos del frente arrecifal no eran estables. Los estudios que de los esquistos ha realizado D. J. W. Piper, de la Dalhousie University, indican que los hundimientos periódicos provocaban un flujo de fango hacia las aguas anaeróbicas y más profundas de la cuenca. Estos flujos barrían todas las huellas superficiales y las madrigueras subsuperficiales dejadas por la fauna de Burgess Shale. Puesto que los animales atrapados por los torrentes de fango morían durante o poco después de su enterramiento, no pudieron dejar nuevas huellas. Esto significa que el tipo de vida de cada especie ha de deducirse a partir del estudio de sus órganos de locomoción y de la comparación con los invertebrados vivos del mismo tipo.

Al propio tiempo, los enterramientos catastróficos en sedimentos anaeróbicos de fino limo en el que los carroñeros no podían sobrevivir, favorecieron en gran manera la conservación de las partes blandas de los animales. Al irse compactando gradualmente el fango y convirtiéndose en dura roca, los cadáveres enterrados resultaron aplastados y las partes blandas se transformaron en delgadas películas de aluminosilicato cálcico. En general, las películas son más bien oscuras, pero determinadas partes de la mayoría de ejemplares se conservan en forma de zonas muy reflectantes.

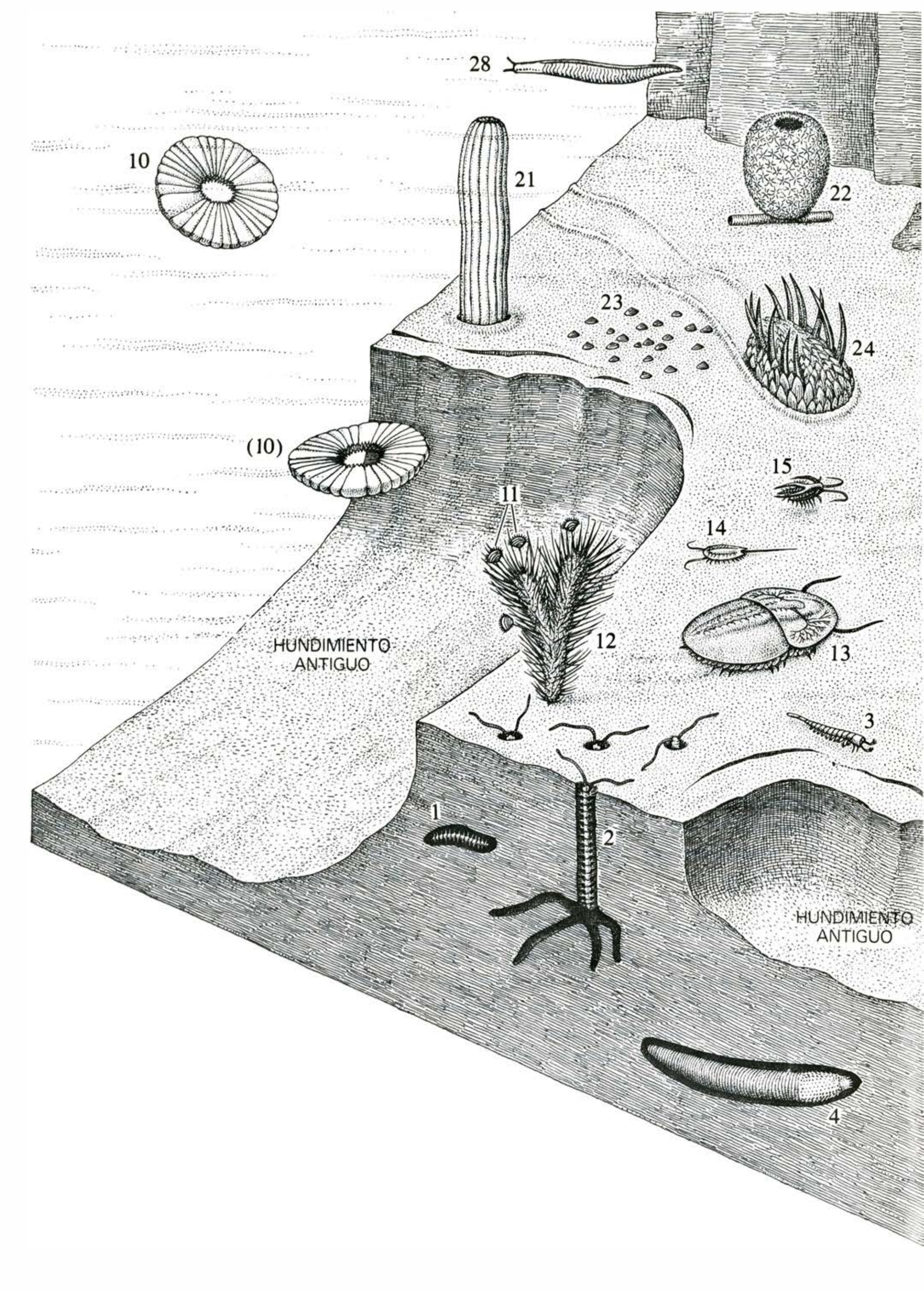
Paradójicamente, y aunque las partes blandas de los animales se hallan maravillosamente conservadas, con frecuencia pueden detectarse indicios de putrefacción después del enterramiento. Mu-

chos ejemplares se hallan asociados a una zona teñida de negro, que es el resultado de haberse infiltrado hasta el fango circundante el contenido corporal del cadáver. En casos extremos, el fósil de un gusano está constituido únicamente por un saco vacío de cutícula porque prácticamente todos los órganos internos del animal se destruyeron debido a la putrefacción. En algunos gusanos, una indicación más sutil de tal desintegración es la separación que existe entre los músculos de la pared del cuerpo y la cutícula.

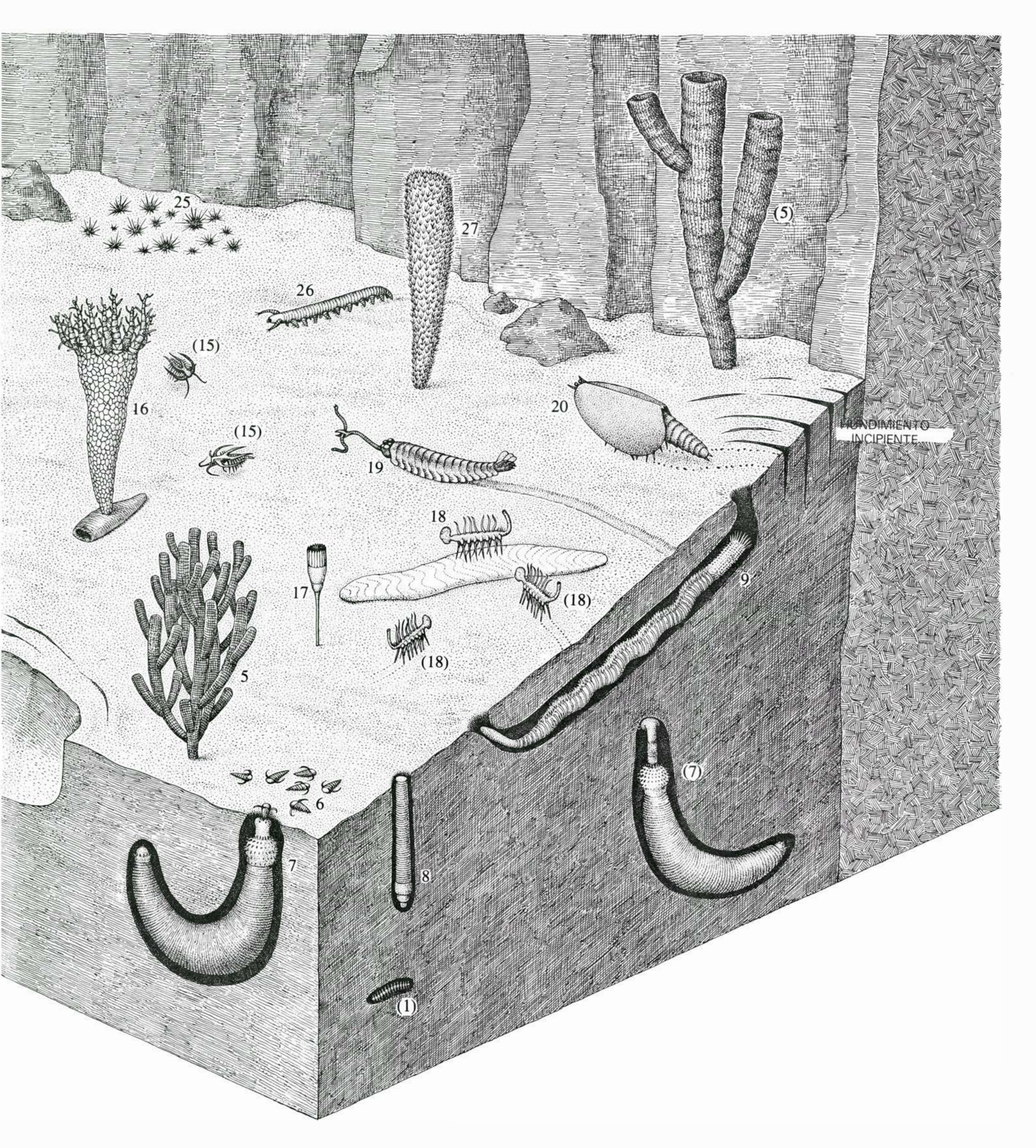
El enterramiento en una corriente de fango tiene otros efectos importantes. Por un lado, muchos de los animales resultan enterrados en todas las posiciones; el yacimiento de esquistos los ha conservado por tanto en diversas orientaciones que revelan mucho más de la anatomía del animal que el simple enterramiento horizontal. Por otro, los sedimentos fluidos que penetraron entre los apéndices de animales tales como artrópodos y poliquetos durante el flujo turbulento de limo quedaron reducidos eventualmente a delgadas capas de esquisto. El empleo cuidadoso del microcincel permite extraer estas finas capas, con lo que quedan al descubierto detalles de la anatomía de un ejemplar que de otro modo no se apreciarían.

La composición de la fauna de Burgess Shale da al traste con la noción convencional de lo que constituye un conjunto típico de animales cámbricos. Los fósiles que se encuentran en la mayoría de localidades del Cámbrico son los exoesqueletos de artrópodos tales como los trilobites, los caparazones de varios miembros del tipo Braquiópodos y de equinodermos tales como los de la clase Eucrinoideos, de exoesqueleto formado por placas. Los animales de este tipo significan únicamente el 20 por ciento de los géneros de invertebrados de Burgess Shale. ¿Está justificado, pues, considerar que el conjunto fósil de Burgess Shale es la norma del Cámbrico, por lo menos en lo que respecta a la fauna de aguas profundas, y considerar que las otras faunas del Cámbrico están sesgadas por la fosilización selectiva de animales con partes duras?

Puesto que Burgess Shale representa un solo ambiente que ha sido congelado durante una fracción de segundo del tiempo geológico, no puede darse a la pregunta una respuesta definitiva. No obstante, varios factores sugieren que la fauna de Burgess Shale no era atípica en los tiempos del Cámbrico medio. La presencia ocasional de especies similares a la de Burgess Shale en otras rocas cámbricas señala la existencia en este perío-



ESCENA SUBACUATICA; ésta es una reconstrucción idealizada del fondo marino donde vivía la fauna de Burgess Shale, con limos cayendo de las paredes del gran arrecife. No se ha intentado presentar a los animales en número proporcional a su abundancia fósil. La fauna se identifica con un número, empezando por la izquierda, abajo; sólo se representa alrededor de una quinta parte de las especies fosilizadas en los esquistos. La mayoría de los animales inmóviles del fondo son esponjas: *Pirania (12)*, dibujada con braquiópodos fijados a sus espículas; *Eiffelia (22)*; *Choia (25)*, gregaria; una especie grácil de *Vauxia (5)*, junto a una especie más robusta arriba a la derecha, y *Chancelloria (27)*. Otros tres animales inmóviles son *Mackenzia (21)*, un celentéreo; *Echmatocrinus (16)*, un crinoideo primitivo, que se ilustra fijado a un tubo vacío de gusano, y *Dinomischus (17)*, una de las especies de Burgess Shale que representan tipos de invertebrados desconocidos hasta el presente. Los animales excavadores con *Peronochaeta (1)*, un gusano poliqueto que se alimentaba de partículas nutritivas que encontraba en el fango; *Burgessochaeta (2)*, otro poliqueto que capturaba el alimento con sus largos



tentáculos; *Ancalagon (4)*, un gusano priapúlido que seguramente es el antepasado de algunos parásitos modernos; *Ottoia (7)*, otro priapúlido, que se ve en el centro alimentándose del molusco *Hyolithes (6)* y a la derecha excavando; *Selkirkia (8)*, un tercer priapúlido, que aquí se ve en una madriguera con el extremo anterior hacia abajo, y *Louisella (9)*, un cuarto priapúlido que vivía en una galería bifurcada y ondulaba su cuerpo para hacer pasar agua oxigenada sobre sus branquias. *Peytoia (10)* es un celentéreo nadador cuya forma es la de una rodaja de piña americana. Los moluscos bentónicos además de *Hyolithes* son *Scenella (23)*, cuyas partes blandas están escondidas bajo conchas en forma de sombrero chino, y *Wiwaxia (24)*, con sus escamas de recubrimiento y sus espinas defensivas, que aquí se ilustra trazando un rastro a través del limo. Entre los numerosos géneros de artrópodos del fondo marino se cuentan *Yohoia (3)*, con sus apéndices prensores distintivos; *Naraoia (13)*, un trilobites atípico que conservaba algunas características larvarias; *Burgessia (14)*, con su larga espina caudal; *Marrella (15)*, que pudo vivir nadando justamente por encima del fondo marino; *Canadaspis (20)*, un crustáceo primitivo, y *Aysheaia (26)*, un animal de patas rechonchas que recuerda a un artrópodo terrestre actual, *Peripatus*. Otros representantes de tipos nuevos que se observan además de *Dinomischus* son *Hallucigenia (18)*, un animal preparándose para alimentarse de un gusano muerto y otros dos acercándose a éste, y *Opabinia (19)*, ilustrado aquí mientras captura un pequeño gusano con su apéndice bifurcado único. Finalmente, nadando solo arriba a la izquierda, se ve a *Pikaia (28)*, el único representante del tipo Cordados (el nuestro) en esta fauna del Cámbrico medio.









do de una fauna de cuerpo blando ampliamente extendida. Además, en algunos conjuntos fósiles del Cámbrico, determinadas especies más bien peculiares pudieron prosperar debido a que el mar tenía un acceso limitado a la zona de deposición. Burgess Shale, en cambio, se encontraba al borde del mar abierto y se hallaba expuesto a la colonización por parte de larvas marinas que llegaban flotando desde otras zonas. Esta circunstancia refuerza la hipótesis de que la fauna de Burgess Shale se acerca a la normal en el Cámbrico.

Debe señalarse al respecto que representantes de algunos invertebrados modernos que casi con seguridad han tenido una larga historia geológica faltan en Burgess Shale. No se encuentra allí ninguna especie del tipo Platelmintos, los gusanos planos, cuyos miembros actuales incluyen las duelas, las tenias y las planarias. Tampoco hay ejemplares de otro tipo de gusanos, los Nemertinos, que comprende los modernos gusanos de trompa, ni tampoco de ninguna especie del tipo Sipunculoideos. Puede que tales gusanos no se hallen representados porque el ambiente del frente arrecifal no fuera adecuado para ellos.

La mayoría de especies que se han encontrado en Burgess Shale pueden situarse en el marco ecológico de una comunidad marina bentónica que prosperó sobre el fondo marino fangoso entre intervalos de hundimientos. El fango mantenía a un grupo activo de invertebrados excavadores, con predominancia de gusanos priapúlidos. Fijadas al fondo del mar y creciendo a distintas alturas, había gran variedad de esponjas que representaban por lo menos 15 géneros; se alimentaban de partículas nutritivas suspendidas en el agua. Rondando activamente sobre la superficie del fondo marino o excavando en el fango en busca de comida había varias especies de artrópodos. Determinados braquiópodos ocupaban un nicho peculiar: se fijaban a las espículas alargadas de una de las esponjas, *Pirania*. Para los braquiópodos las ventajas son evidentes: vivían algo por encima de las aguas turbias del fondo marino y podían capturar partícu-

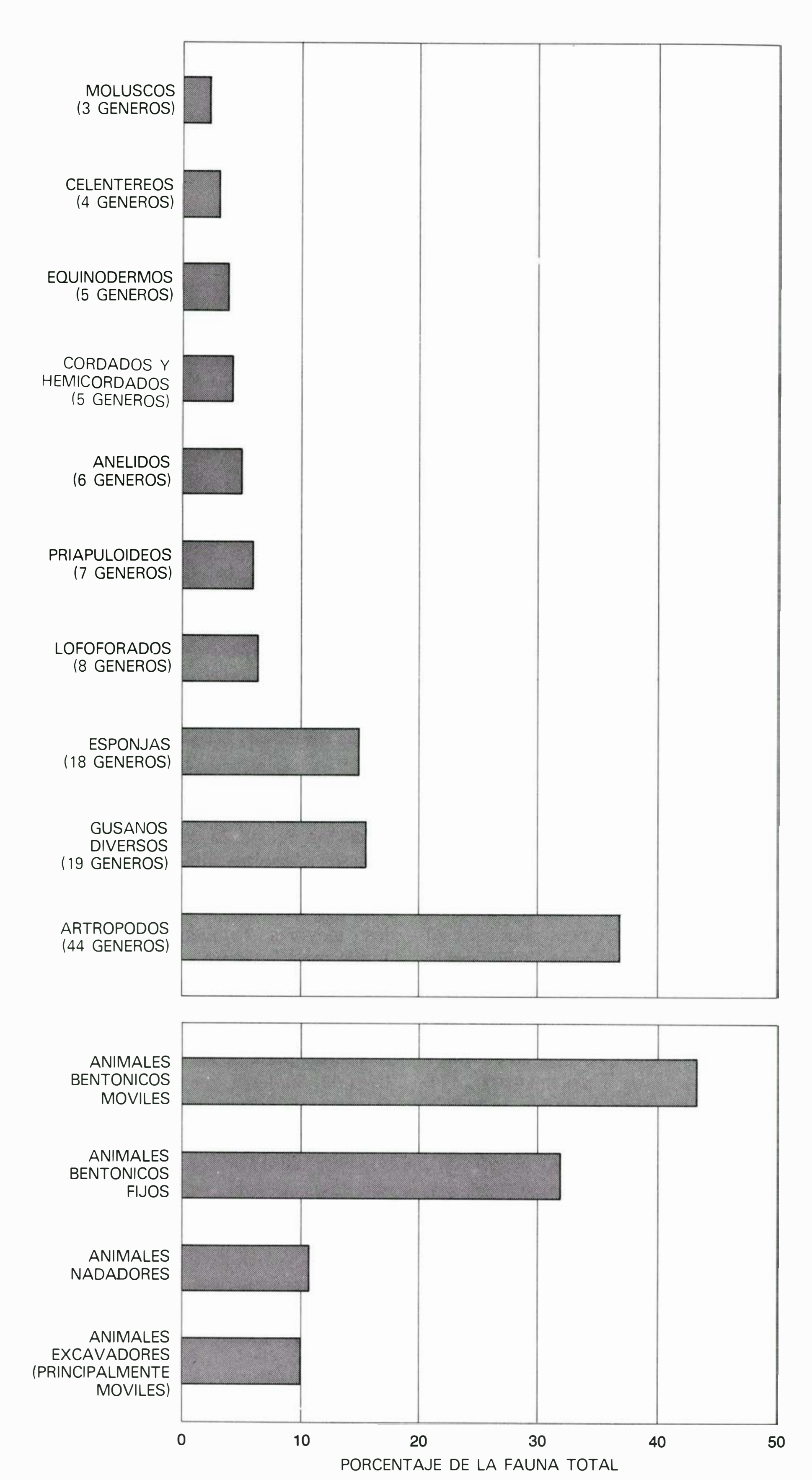
ANIMALES EXCLUSIVOS de la fauna de Burgess Shale. Arriba está *Hallucigenia*, dibujado en la ilustración anterior. El segundo desde arriba es *Nectocaris*, un animal de cuerpo hidrodinámico con aparentes radios de las aletas. Abajo, a la izquierda, está *Amiskwia*, un gusano gelatinoso con aletas prominentes. Abajo, a la derecha, se halla *Dinomischus*, el animal pedunculado cuya reconstrucción se ha hecho también en la ilustración de las dos páginas anteriores. Cada una de estas especies y otras seis o más representan tipos filéticos de invertebrados hasta ahora desconocidos.

las alimenticias como las que nutrían a las esponjas a estos niveles superiores.

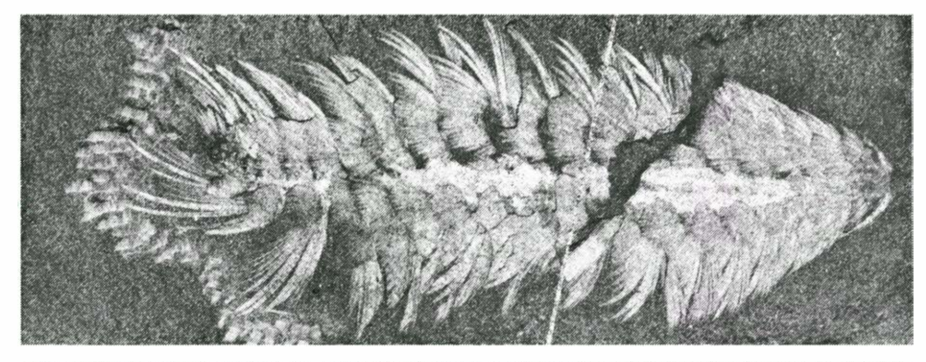
Además de esta comunidad de animales bentónicos fijos y móviles, de superficie y excavadores, varias especies nadadoras vivían en las aguas a lo largo del frente arrecifal. De estos animales quedan sólo retazos exasperantes, en forma de unos raros ejemplares enterrados por casualidad en los sedimentos arrastrados. Los distintos miembros de esta fauna pelágica vivieron probablemente a distintas profundidades; algunos de ellos pudieron haber sido especies arrastradas a la zona frontal del arrecife desde mar abierto.

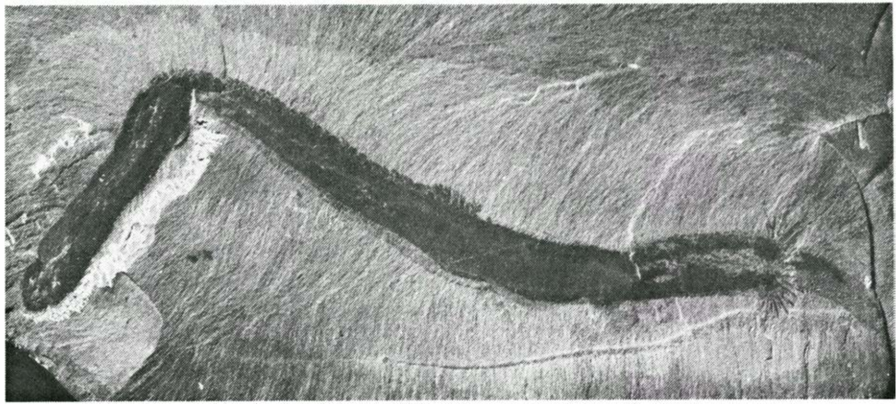
En la mayoría de localidades fósiles del Cámbrico los exoesqueletos mineralizados de los trilobites, los artrópodos más familiares del Paleozoico, son mayoría. Sin embargo, en Burgess Shale los trilobites (con una sola excepción) son comparativamente poco importantes. La excepción es *Olenoides,* de gran significado porque en varios especímenes los apéndices se han conservado con detalle. *Olenoides* poseía un par de antenas delgadas situadas anteriormente y un par de cercos, o estructuras parecidas a antenas, en la parte posterior. Los miembros dispuestos a lo largo de la longitud del animal, hasta 16 en total, tenían todos una construcción similar. La coxa, el segmento mayor y el situado más cerca del cuerpo, presentaba una batería de espinas de aspecto feroz. Unidos a la coxa había dos apéndices: uno era una branquia filamentosa y el otro una pata marchadora. *Olenoides* podía capturar y desmenuzar alimentos blandos, como pequeños gusanos, y pasar los fragmentos hacia su boca. Las antenas delanteras y los cercos traseros suministraban, sin duda, al animal información acerca del alimento y de los depredadores en potencia. El hecho de que las extremidades primitivas de este trilobites sean todas similares constituye un notable contraste con la disposición en muchos artrópodos fósiles y actuales, cuyos apéndices están modificados de forma diversa y especializados.

Alrededor del 40 por ciento de la fauna de Burgess Shale está formada por artrópodos. Tanto en número de especies como en número de especímenes individuales, los representantes de cuerpo blando del tipo sobrepasan en número a los trilobites de caparazón duro. Muchos de estos "no trilobites" tienen los apéndices preservados con un detalle notable; algunos de ellos debieron haber sido depredadores y carroñeros eficaces. El más abundante es *Marrella*, un artrópodo con una cabeza cuneiforme que



LOS GENEROS DE BURGESS SHALE son hasta ahora 119. El porcentaje del total que corresponde a los distintos tipos se indica en la parte superior de este histograma. Casi el 40 por cien del total está constituido por artrópodos; sólo 14 de los 44 géneros de artrópodos son trilobites. Los gusanos que no son priapuloideos ni anélidos (19 géneros) y las esponjas (18 géneros) constituyen otro 30 por cien del total; los moluscos son los menos representados. En términos de hábitat, como indican los histogramas inferiores, más del 40 por cien de los animales de Burgess Shale se desplazaban sobre el fondo, y más del 30 por cien se hallaban fijados en el fango. La mayoría de animales excavadores se movían también libremente; las especies nadadoras sobrepasaban ligeramente en número a los animales excavadores.





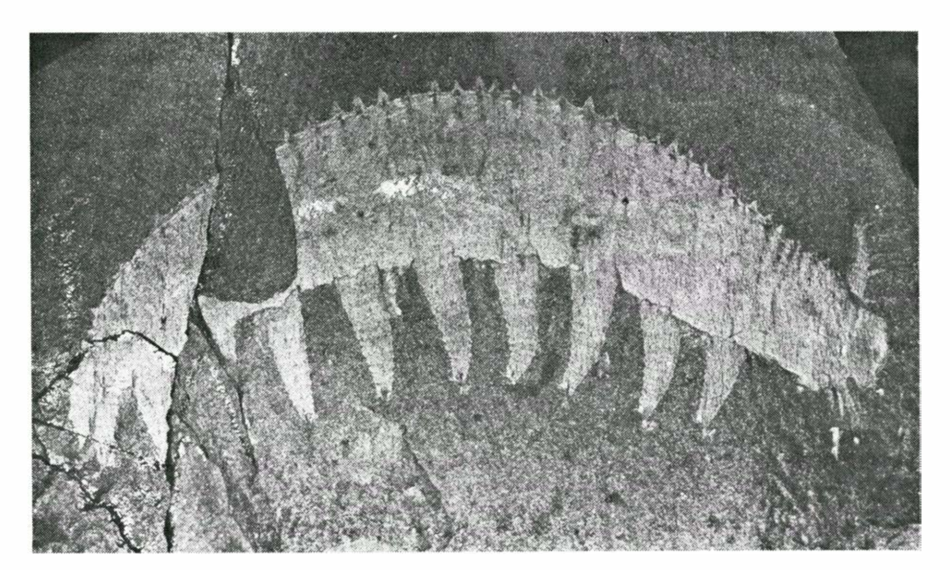
DOS DE LOS TIPOS DE GUSANOS de la fauna de Burgess Shale son los familiares Anélidos y los menos comunes Priapuloideos. Un anélido típico de Burgess Shale es *Canadia*, un gusano poliqueto (*arriba*); sus sedas, haces de finas cerditas que era órganos de locomoción, se han preservado en detalle. *Louisella* (*abajo*) es un típico gusano priapúlido, que se ha reconstruido en una ilustración anterior.

portaba dos pares de largas espinas en forma de cuernos que se curvaban hacia atrás. *Marrella* tentaba el ambiente del fondo marino con un par de antenas y dirigía partículas alimenticias hacia su boca con unos apéndices plumosos adyacentes. Sus veinte o más apéndices laterales eran articulados, y de cada uno de ellos se ramificaba una branquia filamentosa.

El no trilobites que sigue en abundancia es *Canadaspis.* Todo su cuerpo, a excepción de la parte posterior, se hallaba escondido por un caparazón doble. La remoción cuidadosa de esta cubierta revela los apéndices subyacentes, notablemente parecidos a los de algunos crustáceos actuales. Otro artrópodo era uno de los mayores depredadores sobre el fondo marino; se trata de *Sidneyia*, animal cuyas extremidades distintivas recuerdan a las de la cacerola de las Molucas actual, *Limulus.* Ha sido posible identificar parte de los contenidos digestivos de *Sidneyia*, que han resultado ser fragmentos de caparazones de braquiópodos, lo que prueba que este artrópodo podía triturar presas de cuerpo duro.

Sidneyia no era el mayor artrópodo de Burgess Shale. Los esquistos contienen las impresiones de grandes patas aisladas; si se mantiene la proporción, pueden haber pertenecido a un animal de hasta un metro de longitud. A estos fósiles se les ha asignado el nombre de *Anomalocaris*, y posiblemente representan los mayores invertebrados del Cámbrico.

El más interesante de todos los artrópodos de Burgess Shale es *Aysheaia*, un animal con un cuerpo rechoncho y extremidades regordetas. Cuando Walcott publicó por primera vez una fotografía de este fósil hace medio siglo, varios zoólogos le escribieron para indicarle que este invertebrado del Cámbrico medio se parecía mucho a *Peripatus*, un animal con ocho géneros emparentados (que comprenden dos familias) en la pequeña clase Onicóforos, dentro del tipo Artrópodos. *Peripatus* es un animal terrestre, y *Aysheaia* era una forma marina; no obstante, *Aysheaia* representa seguramente el tipo de antepasado que pudo haber dado origen a artrópodos actuales como Miriápodos e Insectos.



UN ARTROPODO ANCESTRAL, con un parecido sorprendente al onicóforo actual *Peripatus*, es este invertebrado notablemente bien conservado, *Aysheaia*. Los artrópodos cámbricos como *Aysheaia* pudieron haber sido los antepasados originarios de miembros actuales de este tipo, los Miriápodos y los Insectos.

Aunque algunos de los artrópodos no trilobites de los esquistos de Burgess Shale, como *Aysheaia*, recuerdan formas posteriores, la mayoría de ellos no pueden ser colocados en ningún grupo reconocido. No tienen parientes claros entre las demás especies de Burgess Shale ni entre los artrópodos de épocas posteriores. Debido a que muestran una variación extraordinariamente amplia de características anatómicas, lo que indica un elevado grado de especialización, constituyen la prueba de una radiación adaptativa de los artrópodos en el Cámbrico, radiación hasta ahora insospechada. Parece que las numerosas poblaciones que surgieron durante este período de rápida evolución fracasaron en su mayor parte. Es de interés señalar que los animales que iban a resultar dominantes en la historia geológica posterior tienen por lo general una posición menor en las distintas faunas del Cámbrico; un observador hipotético se hubiera visto en apuros a la hora de predecir qué grupos tenían la flexibilidad necesaria para el éxito biológico a largo plazo.

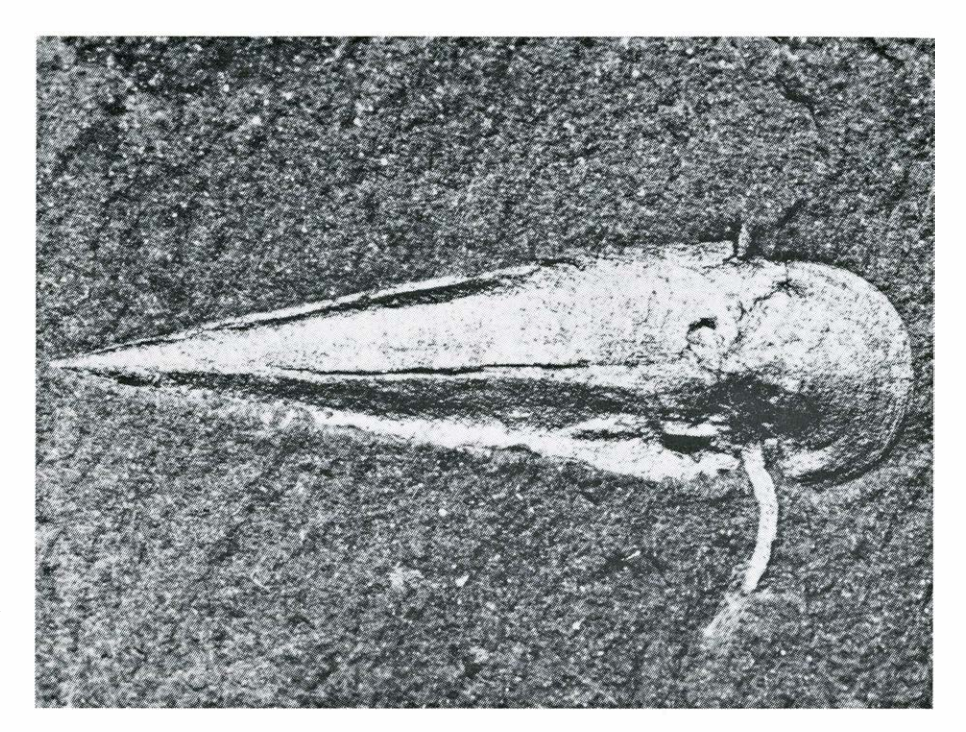
De los animales de Burgess Shale distintos de los artrópodos, son particularmente notables los representantes de seis tipos. Entre los Equinodermos, se pensó antaño que la clase Holoturioideos (el grupo que incluye las holoturias o cohombros de mar) estaba ampliamente representada. Hoy sólo se clasifica así a

un animal, *Eldonia*. A diferencia de la gran mayoría de especies en su clase, *Eldonia* tiene un cuerpo medusoide y un par de tentáculos orales. Estos animales debían probablemente nadar a través del agua en cardúmenes, utilizando sus tentáculos para capturar el alimento. Otro equinodermo de Burgess Shale, el lirio de mar *Echmatocrinus*, es el crinoideo más antiguo del registro fósil; como podía esperarse, presenta numerosos rasgos primitivos.

Las especies del tipo Celentéreos figuran entre los metazoos más primitivos. Por ejemplo, en el conjunto fósil del Precámbrico tardío de Ediacara Hills predominan los celentéreos. En contraste, los distintos celentéreos de Burgess Shale, algunos similares a medusas y otros a plumas de mar, parecen haber desempeñado un papel bastante limitado en la comunidad. Por otro lado, las esponjas de Burgess Shale, los más primitivos de todos los animales que allí se encuentran, eran miembros importantes de la comunidad. Eran abundantes y de formas variadas; algunas especies crecían sobre el fondo formando matas.

Los diferentes "gusanos" de Burgess Shale fueron asignados principalmente por Walcott al tipo Anélidos en general y a la clase Poliquetos en particular. Hoy se sabe que muchos de ellos pertenecen a otros tipos. No obstante, entre los gusanos poliquetos se encuentran algunos de los ejemplos más espectaculares de conservación de estructuras blandas: las sedas, o haces de finas cerditas que constituían los órganos de locomoción de estos animales, se han conservado particularmente bien en forma de películas brillantes y reflectantes en los esquistos. Uno de los poliquetos, *Canadia*, aparentemente no vivía en una madriguera, sino que pasaba gran parte de su tiempo nadando pegado al fondo marino. Otro, *Burgessochaeta*, era probablemente un excavador más típico, refugiándose en el fondo fangoso y buscando su alimento alrededor de la entrada de su madriguera con sus largos tentáculos.

Hoy día, el tipo Priapuloideos es de interés sólo para un puñado de especialistas. Estos gusanos, sin embargo, constituían un grupo importante en el Cámbrico, y dos priapúlidos presentes en Burgess Shale son particularmente dignos de mención. Uno de ellos, *Ottoia*, es el más abundante del grupo. Ha llegado hasta nosotros con tanto detalle que los músculos son claramente visibles y puede analizarse el contenido digestivo de algunos especímenes. *Ottoia* se alimentaba de dos tipos de animales: braquiópodos e hiolítidos. Los Hiolítidos,



UN MOLUSCO REPRESENTATIVO, *Hyolithes*, poseía una concha cónica rematada por un opérculo protector. Uno de los gusanos excavadores, *Ottoia*, hacía presa en estos moluscos pero no podía abrir la concha. El gusano se limitaba a digerir las partes blandas de *Hyolithes* y excretaba su concha protectora.



UN PROBABLE MOLUSCO, *Wiwaxia*, con su recubrimiento de grandes escamas y su dotación de largas espinas protectoras, fue clasificado primero entre los gusanos poliquetos de Burgess Shale. Sin embargo, su órgano alimentario raspador, similar a la rádula, sugiere su pertenencia al tipo Moluscos.

posiblemente miembros del tipo Moluscos, poseían una concha cónica coronada por un opérculo o tapadera protectora cuando el animal se retiraba por completo en su interior. Los dientes de *Ottoia* no eran suficientemente fuertes para romper la concha, de modo que los hiolítidos eran tragados enteros y sus partes blandas digeridas mientras las conchas pasaban incólumes a través del tubo digestivo del priapúlido. Estos mariscos no eran el único alimento de *Ottoia*. Un ejemplar único contiene en su tubo digestivo los restos de otro gusano de la misma especie, lo que demuestra que (como en algunos gusanos priapúlidos actuales) *Ottoia* podía ser caníbal.

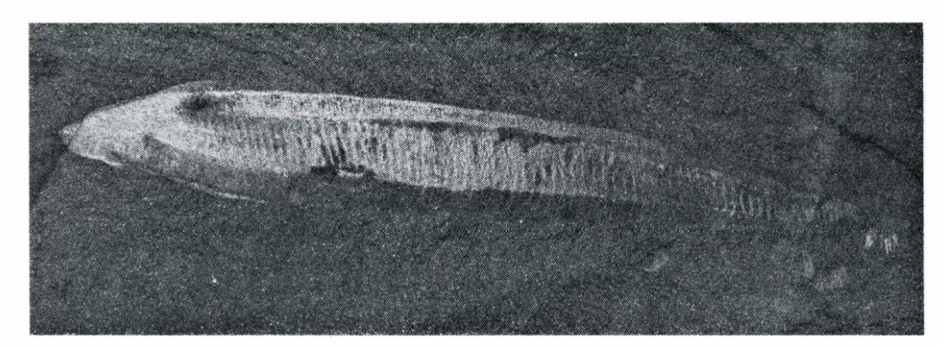
Los parasitólogos han demostrado un interés considerable por otro priapúlido de Burgess Shale. Se trata de *Ancalagon*, que pudiera ser el antepasado del grupo actual de gusanos de cabeza espinosa, los Acantocéfalos, que parecen haber sido parásitos durante millones de años. Estos gusanos parásitos carecen de tubo digestivo y absorben el alimento a través de la pared corporal mientras están alojados en el intestino de su hospedante. Si se invierte hipotéticamente la evolución y se vuelve a dotar a estos gusanos con los órganos necesarios para una existencia libre, el animal reconstruido es notablemente parecido a *Ancalagon*.

Otros dos supuestos gusanos, que antes se consideró que eran poliquetos, son *Wiwaxia* y *Pikaia*. El cuerpo de *Wiwaxia* estaba cubierto con grandes escamas. Unas espinas largas que se curvaban hacia arriba y hacia afuera a lo largo del dorso del animal eran, evidentemente, una protección contra los depredadores; que las espinas eran realmente protectoras lo indica el hecho de que en algunos ejemplares de *Wiwaxia* han sido arrancadas. Este animal vivía en el fondo marino y se alimentaba rayendo fragmentos de comida con un órgano raspador. Esta raspa se parece a la rádula, o lengua dentada córnea, de algunos moluscos actuales. ¿Es *Wiwaxia* un molusco primitivo? Si es así, los detalles de su anatomía notablemente conservada aportarán nueva luz a la evolución temprana de este tipo de invertebrados tan próspero.

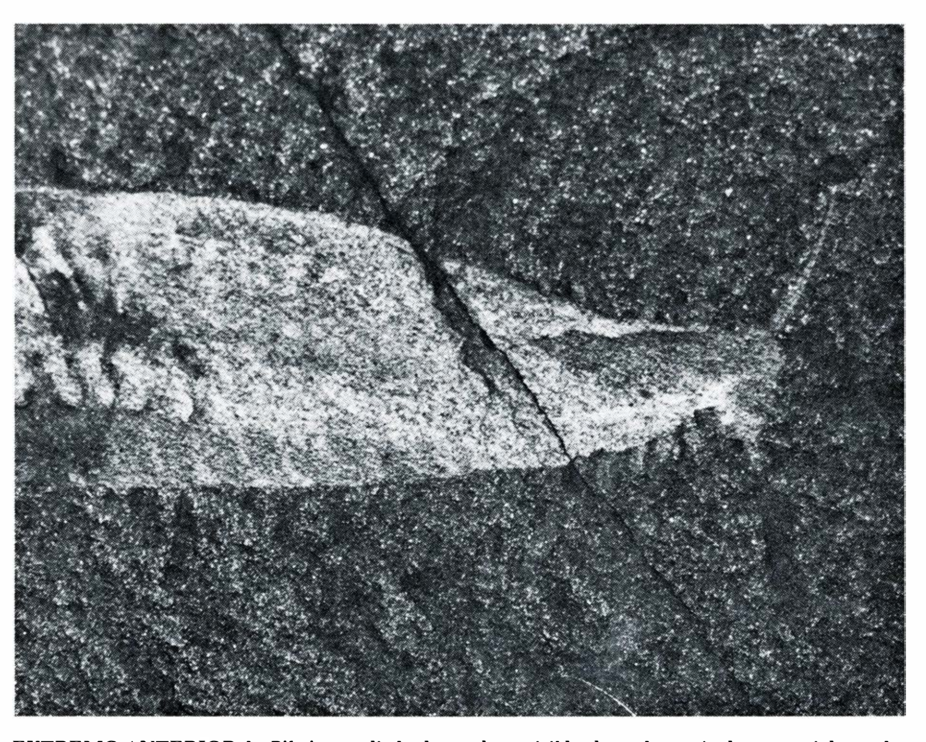
¿Y qué hay de *Pikaia*, considerado en un principio como un gusano poliqueto? Unos 30 especímenes bien conservados muestran una varilla prominente a lo largo del dorso del animal que parece ser un notocordio, el órgano de refuerzo semicartilaginoso que da su nombre al tipo Cordados. Además de esta característica anatómica clave, los bloques musculares de *Pikaia* forman una estructura en zigzag que es comparable a la musculatura del primitivo cordado actual *Branchiostoma* y de los peces. Aunque *Pikaia* difiere de *Branchiostoma* en varios aspectos importantes, la conclusión de que no es un gusano sino un cordado parece ineludible. La maravillosa conservación de este organismo del Cámbrico medio hace de él un hito en la historia del tipo filético al que pertenecen todos los vertebrados, incluido el hombre. Existen posibles casos de cordados aún más primitivos procedentes de formaciones del Cámbrico inferior de California y Vermont, pero ninguno es tan rico en detalles.

Quizá el problema más intrigante que plantea la fauna de Burgess Shale está constituido por los 10 o más géneros de invertebrados que hasta ahora han desafiado todos los esfuerzos que se han hecho para relacionarlos con tipos conocidos. Parecen ser los únicos representantes identificados de tipos filéticos cuya existencia no se había sospechado anteriormente. Sus orígenes deben hallarse en la oscuridad del Precámbrico, donde empezó la diversificación inicial de los Metazoos. La peculiaridad de estos animales nuevos queda ejemplificada por el adecuadamente llamado *Hallucigenia*.

Este animal se impulsaba sobre el fondo marino mediante siete pares de espinas muy agudas y en forma de zancos. De la cara superior del cuerpo del animal surgían siete tentáculos, en el extremo de cada uno de los cuales había un par de puntas reforzadas. ¿Capturaban alimento estos tentáculos? Si lo hacían, ¿actuaba cada tentáculo como una boca individual con una conexión directa con el canal alimentario del animal? Hay más preguntas que respuestas, pero una clave valiosa para el comportamiento del animal se halla preservada en un espécimen de una colección de la Universidad de Harvard. En él pueden verse más de 15 individuos de *Halluci-*



UNA VARILLA DE REFUERZO, o notocordio, corre a lo largo de parte del dorso del primitivo cordado *Pikaia*. La cabeza del animal, que se ve con más detalle en la ilustración de abajo, se halla a la derecha. La estructura de la musculatura se parece a la de los peces y a la del cordado primitivo actual *Branchiostoma*. En una ilustración anterior aparece una reconstrucción de este cordado nadador.



EXTREMO ANTERIOR de *Pikaia*, ampliado, lo que hace visible el par de tentáculos sensoriales, y detrás de ellos una corta fila de pequeños apéndices. En formaciones del Cámbrico temprano se conservan restos de posibles cordados, pero ninguno de ellos es comparable con *Pikaia* en cuanto al detalle.

genia asociados a un gusano grande. No parece haber dudas de que, habiendo detectado el cadáver del gusano, estos raros animales se congregaron para alimentarse de sus restos.

Comparado con *Hallucigenia*, un segundo animal único, *Opabinia*, parece casi ortodoxo. Tenía cinco ojos dispuestos a través de la cabeza, de modo que probablemente podía evitar fácilmente a los depredadores mientras nadaba cerca del fondo, dirigiéndose gracias a una aleta caudal vertical. *Opabinia* se alimentaba capturando presas mediante un órgano prensor que se proyectaba hacia adelante. En estos invertebrados insólitos se evidencian distintas soluciones a problemas de diseño funcional. Así, para un gusano con una cavidad del cuerpo llena de líquido, un problema es que la contracción muscular en una parte del cuerpo distorsionará la forma del resto del mismo. En los gusanos anélidos el problema se ha resuelto dividiendo la cavidad del cuerpo en una serie de compartimentos estancos. *Banffia*, un gusano único de Burgess Shale, desarrolló una solución alternativa. La mitad anterior más rígida de su cuerpo estaba separada de la mitad posterior más sacciforme por una constricción prominente en el punto medio. La constricción parece haber amortiguado las fluctuaciones hidrostáticas desencadenadas por los músculos locomotores de la mitad anterior del animal, minimizando así la distorsión de su mitad posterior no rígida.

Algunos representantes de tipos nuevos han llegado hasta nosotros por docenas. Otros, en particular los habitantes nadadores de los niveles de agua superiores que raramente debían ser atrapados por los hundimientos de fango, son bastante raros. Uno de tales animales es el gusano *Amiskwia*; a juzgar por sus aletas prominentes, con probabilidad se trataba de un ágil nadador. Otro animal, *Nectocaris*, un depredador de natación rápida, tenía ojos enormes y evidentemente impulsaba su cuerpo hidrodinámico mediante rápidos movimientos laterales del mismo. Prominentes aletas dorsales y ventrales, endurecidas por numerosos radios, ayudaban a mantenerlo estable mientras éste nadaba.

Los conodontos, o "dientes cónicos", son fósiles enigmáticos que parecen dientes diminutos; se encuentran en formaciones cuya edad va desde el Precámbrico más tardío hasta el Triásico, es decir, a lo largo de casi 400 millones de años. Aunque parecen dientes no pueden haber funcionado como tales porque muestran señales de desgaste. Durante mucho tiempo han sido preguntas sin respuesta qué animal de cuerpo blando tenía conodontos, y con qué fin. Otro raro invertebrado pelágico preservado en Burgess Shale, *Odontogriphus*, pudiera ser este animal. El aparato alimentario tentacular del animal, otro representante único de un tipo hasta ahora desconocido, posee una dotación de minúsculos objetos cónicos que parecen ser conodontos. Puesto que los conodontos no pudieron haber funcionado como dientes, se ha sugerido la hipótesis de que eran algún tipo de soporte para los tentáculos alimentarios. ¿Fue el aparato alimentario de *Odontogriphus*, y de animales como él, el origen de los conodontos tan abundantemente distribuidos por todo el registro fósil del Paleozoico y del Mesozoico temprano? Posiblemente.

Cuando se sepa más de la fauna de Burgess Shale, el cuadro de la vida en el Cámbrico ganará una nueva perspectiva, en particular con respecto a la evolución explosiva de los Metazoos. Por ejemplo, la gama de artrópodos, con sus agrupaciones distintivas y diferentes de características anatómicas, es ya tan amplia que un solo tipo filético parece demasiado pequeño para contenerlos a todos. La radiación adaptativa de los invertebrados del Cámbrico puede considerarse como la respuesta inicial a la disponibilidad de una gran variedad de nichos ecológicos marinos. De aquí que muchos animales cámbricos semejen experimentos pioneros realizados por varios grupos de metazoos, destinados a ser suplantados a su debido tiempo por organismos mejor adaptados. La tendencia después de la radiación del Cámbrico parece ser el éxito y el enriquecimiento en el número de especies de relativamente pocos grupos, a expensas de la extinción de muchos otros.

Una posibilidad adicional viene sugerida por la propia fauna de Burgess Shale. A algunos grupos de gran entidad en el Cámbrico, como los gusanos Priapuloideos, les pudo haber ido mal frente a competidores posteriores, y sólo se libraron de la extinción migrando a nichos marginales que no eran atractivos o no eran accesibles a los demás metazoos. Una de tales manifestaciones de movimiento hacia un nicho marginal es la reducción progresiva del tamaño del cuerpo. Esta miniaturización puede muy bien ser la razón por la que algunos priapúlidos lograron sobrevivir. Una vía de huida alternativa es hacerse parásito; los priapúlidos que parecen haber dado origen a los acantocéfalos parásitos pueden ser un ejemplo de ello. En cualquier caso, la fauna de Burgess Shale permite una ojeada maravillosa a la evolución en acción durante este breve intervalo del Cámbrico medio.

Frege

La publicación de Begriffsschrift en 1879 supone el nacimiento de la lógica matemática. Su objetivo central: fundamentar la aritmética en el pensamiento puro

Javier de Lorenzo

El año 1879 se viene considerando como una de las fechas clave en la lógica matemática: la de su nacimiento. Es el año en que se publica *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens,* escrito por el matemático Gottlob Frege, y que traduzco por "Ideografía, un lenguaje de fórmulas para el pensamiento puro modelado en el lenguaje de la aritmética" y que citaré, en adelante, por *Bs.* Frege nació en 1848 en Wismar y murió en 1925 en Bad Kleinen. Estudió y se habilitó en Göttingen, en 1873, pasando a ser privat-dozent en la Universidad de Jena desde 1874 hasta 1896 y profesor honorario desde ese año hasta 1917. Trabajó en soledad y sin muchos reconocimientos en vida.

Según el secretario de la Universidad de Jena "su actividad académica carecía de interés para la Universidad", lo que vendría corroborado por la afirmación de Carnap al contar que, en 1913, sólo asistían al curso de Frege, con él, otros dos alumnos, uno de ellos comandante retirado que deseaba estar al tanto de las "últimas" novedades. Hecho bastante más normal de lo que parece, aunque algunos comentaristas quieran descargar su conciencia hablando del genio incomprendido, etc. Y si digo "normal", basta mencionar en el entorno de esas mismas fechas a Kummer, quien jamás llegó a profesor universitario; a Cantor, que no pudo lograr su aspiración de llegar a Berlín; a Kronecker, que si dio clases en la Universidad, las dio no por catedrático o profesor ordinario, sino por académico; a Dedekind, que se mantuvo en su Escuela Técnica de Brunschwich, alejado del "académico y universitario ruido".

Frege estudió en Göttingen. Y Göttingen ha sido el centro de la matemática durante el siglo XIX y primeros años del XX. Desde Gauss hasta Hilbert pasando por Riemann y Weiertrass. Naturalmente, el estudiante de Göttingen se encontraba en el centro de los motivos y de los problemas vivos de la matemática. Y Frege no pudo sustraerse a ellos. Especialmente a uno: fundamentar la aritmética y aclarar de una vez para siempre la naturaleza de los números naturales. Es el objetivo de su vida, que se condensa en lo que viene estimándose *programa logicista* en la fundamentación de la matemática: reducir la aritmética a la lógica, es decir, derivar los conceptos de la aritmética de conceptos lógicos y deducir los principios aritméticos de los principios lógicos. La aritmética, el número natural, como elementos del pensamiento puro, sin intervención de imágenes, percepción, objetos materiales.

Frege creyó por un momento que había conseguido su objetivo hacia 1902, pero las antinomias dejaron al descubierto que no había analizado bastante, que la construcción en la que había empeñado su vida se venía abajo. En su diario escribirá un año antes de su muerte: "Mis esfuerzos por aclarar lo que sean los números han conducido a un completo fracaso". Y posteriormente: "Me he visto obligado a abandonar la opinión de que la aritmética sea una rama de la lógica y, por tanto, que todo en la aritmética puede ser probado lógicamente".

He dicho Göttingen, y no sólo en Göttingen se hacía matemática. En otros lugares he estudiado el proceso que sufre la matemática en los entornos de 1827. Nuevos temas –geometría proyectiva, geometría diferencial, análisis...– y, fundamentalmente, nuevo enfoque en el hacer matemático. Enfoque que se centra en hallar la razón y no ir de lo particular a lo general, sino de lo general a lo particular (Abel). Un hallar la razón interna a la matemática y no supeditado a las ciencias de la naturaleza, al empirismo utilitario. Un hallar la razón que, por no poseer criterio extrínseco, sólo podrá realizarse apoyándose en criterios estrictamente racionales: realizar las demostraciones con todo rigor, para lo cual es preciso un previo análisis de los conceptos que entran en juego. A pesar de las nuevas geometrías, la representación geométrica se convierte, de ayuda, en constante peligro. Weiertrass construye una curva que, continua en todos sus puntos, carece de tangente en cada uno de ellos, y da paso a toda una familia de curvas teratológicas. Desde sus comienzos, el cálculo hace uso de la noción de número real, del continuo, que no está definido. El único campo aparentemente seguro es la aritmética. Y sobre él se pretenderá fundamentar el análisis. Es lo que vino a denominarse "proceso de aritmetización del análisis". Proceso que culmina en 1872 con la caracterización de los números reales por Weiertrass, Cantor, Méray, Dedekind simultánea, independientemente y con procesos diferentes. Culminación que supone un cambio cualitativo, una ruptura epistemológica radical en el interior de la matemática y un inicio de un nuevo tipo de hacer matemático. Nuevo tipo que, en los terrenos de la geometría, tiene su paralelo con el programa de Erlangen enunciado por Klein el mismo año de 1872.

Suponen, los años que entornan 1875, un cambio en el estatuto del hacer matemático. Ya no se trata de manejar elementos y correspondencias de elemento a elemento, o magnitudes que impliquen una medida y un contar, sino de manejar sistemas, clases o conjuntos dados en acto y establecer aplicaciones y relaciones entre ellos, así como entre los elementos y los sistemas a los cuales pertenecen. Conjuntos o sistemas que pueden poseer infinitos elementos. Para manejarlos, surgen, como conceptos básicos, los de función o aplicación con sus diversos tipos y la relación de equivalencia con su conjunto cociente asociado. Son los que permitirán probar teoremas fundamentales de la continuidad de funciones de variable real, tema que surge nuevo o, en manos de Dedekind, los que van a permitir la caracterización no sólo de los números reales mediante su proceso de cortaduras sino

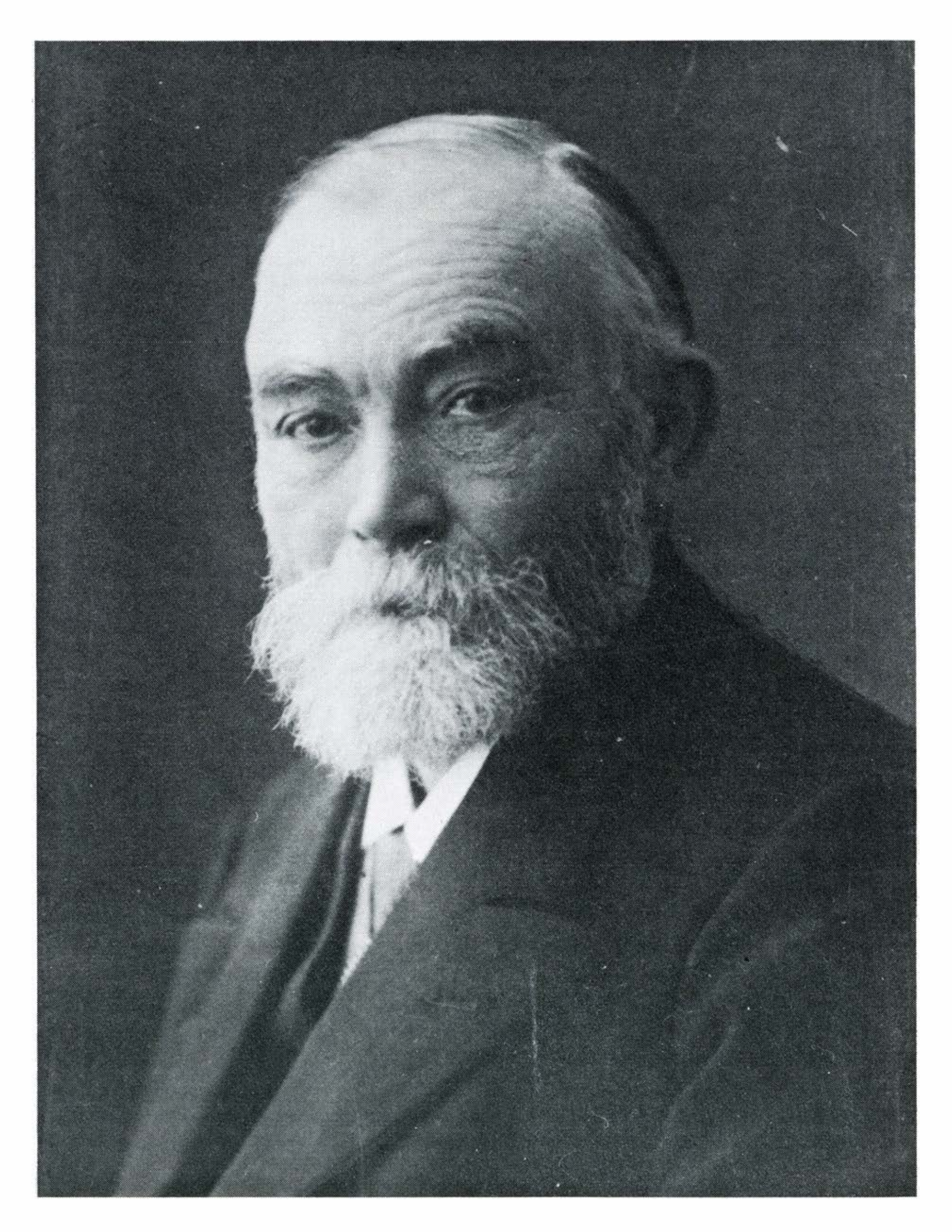
la creación del álgebra moderna tras la elaboración del concepto de *ideal*, con la convicción subyacente de que esos sistemas o conjuntos no pueden ser reuniones arbitrarias de elementos cualesquiera, sino de elementos con una determinada estructura, que hay que caracterizar mediante "leyes formales", estrictamente lógicas, entendiendo por tales las que emanan del pensamiento puro y no las procedentes de la experiencia sensible o psicológica.

Como objetos de la "matemática moderna" se estimarán los conjuntos y las aplicaciones, como escribirá Clifford en 1872, y posteriormente recordará Frege en 1884. Y si éstos son los conceptos básicos de la matemática, surgidos tras los intentos de fundamentar en la aritmética el resto de las disciplinas matemáticas, se tiene que esa misma aritmética carece, a su vez, de fundamento. No parece bastar una convicción, como la que después enunciará Kronecker, de que los números naturales son obra de Dios y el resto de los hombres. Tampoco que existen unas leyes formales aritméticas que deben mantenerse en toda ampliación de sistemas. Aunque es este tipo de convicción el que predomina en el ambiente matemático que rodea este período de creación y que plasmó Hankel en 1867 con su principio de permanencia de leyes formales. Convicción de que, a partir de la aritmética, lo único que importa es el signo material y los postulados que regulan tales signos. Convicción de mero inscripcionismo que confunde signo con lo representado por dicho signo y que, por ello, muestra una debilidad conceptual absoluta desde su inicio.

A pesar de esta convicción, más o menos difundida en el trabajador matemático, y ya desde la nueva ruptura conceptual, se pretende fundamentar la propia aritmética y no sólo tomarla como mero modelo. Fundamentación, por supuesto, no psicológica o imaginativa, sino lógica. Fundamentación que no sólo indique qué sea un número natural, sino que regule lógicamente el tipo de razonamiento específico de la aritmética, el principio de inducción completa, si es que ello es posible. Y a esta labor se verán llevados, por unos u otros caminos, matemáticos como Cantor, Dedekind, Schroder, Peano, Peirce, Frege... Desde otro enfoque, también se ligará al tema Husserl, que prefiere abandonar en 1884 la ayudantía con Weiertrass y su tesis doctoral sobre el cálculo de variaciones para, desde el terreno fenomenológico, fundamentar la aritmética, realizar su *Filosofía de la aritmética.*

Por supuesto, los nombres que he citado no son los únicos, ni el intento de fundamentar la aritmética el tema único de estos matemáticos, salvo en Frege. Precisamente no cabe olvidar que, a partir de Boole, De Morgan, Hamilton... se había producido en los terrenos de lo que considerar lógica una cierta renovación. Especialmente, George Boole, desde 1847, y partiendo del hacer matemático, había realizado la construcción de un álgebra lógica. Iniciaba así un proceso de algebrización de la lógica, de aplicar el álgebra para fundamentar la lógica. Boole parte de las nociones de clase, elemento de clase y operaciones con clases. El enfoque estriba en que las leyes del pensamiento, las leyes de la lógica, deben ser del mismo tipo que las que gobiernan el álgebra; es decir, la validez de los procesos del álgebra no depende de la interpretación de los signos, sino de las leyes de combinación de los mismos. Con estas ideas, Boole algebriza la lógica obteniendo un sistema algebraico que es, en términos actuales, un retículo booleano. Dicho retículo, como álgebra lógica o álgebra simbólica, presenta dos interpretaciones: por un lado, un álgebra de clases; por otro, un álgebra proposicional. En el fondo, una misma teoría formal con dos interpretaciones diferentes; según el modo de lectura o la interpretación que se asigne a las variables y a los operadores.

En su construcción, Boole utiliza la aritmética no sólo como modelo, sino que los mismos signos aritméticos son empleados en una nueva acepción; así, la suma y el producto será ahora suma y



GOTTLOB FREGE nació en 1848 en Wismar y murió en 1925 en Bad Kleinen. Estudió y se habilitó en Göttingen, en 1873, pasando a ser privat-docent en la Universidad de Jena desde 1874 hasta 1896 y profesor honorario desde ese año hasta 1917. Según el secretario de dicho centro "su actividad académica carecía de interés para la Universidad". Actualmente se le estima como el lógico más grande de todos los tiempos, sólo equiparable a la figura de Aristóteles quien fundó la lógica en sus *Analíticos.*

producto lógicos; una ecuación como $x = 1$ será ahora “ ‘x’ es verdadero” en interpretación proposicional, la clase universal en interpretación de álgebra de clases...

Y en la misma orientación de Boole de fundamentar la lógica en la matemática, de construirla como una estructura simbólica formal con ulteriores interpretaciones, se encontrarán Jevons, Schroder... Este último publica en 1877 un breve folleto en el que, apoyándose en la línea booleana, trata de perfeccionarla. Posteriormente, consigue formular un sistema axiomático, hoy clásico, caracterizador de la noción de retículo, con su principio de dualidad. Es línea de decisiva importancia en el ulterior desarrollo de la lógica matemática, por la influencia no sólo de Peano, sino porque fue retomada por Skolem, la escuela polaca y, fundamentalmente, por Tarski y quienes siguen la tendencia semántica, que acaba reflejándose en la teoría de modelos y en los intentos de algebrizar los sistemas lógicos de primer orden mediante las álgebras cilíndricas de Tarski, las poliádicas de Halmos. Igualmente, y con parecida tendencia, pero ahora muy condicionada por la aritmética, se manifiesta esta tendencia en la teoría de funciones recursivas, o en la construcción de modelos no canónicos del análisis a partir de ultrafiltros...

Habría que agregar la figura del norteamericano Peirce. Pero sus trabajos, bastante dispersos, en su aislamiento propio, sólo cobrarán todo su valor desde la perspectiva de los logros posteriores.

Admitido por los matemáticos, a partir de 1872, que todos los conceptos de la matemática pueden reducirse a los de la aritmética y los de ésta a los números naturales, Frege adopta sobre sí la tarea de derivar estos últimos por medios estrictamente lógicos. Con ello lograría establecer que toda la matemática es reducible a la lógica. Para esta labor tiene que cumplir dos objetivos: (1) precisar qué entiende por lógica y enumerar los conceptos lógicos con los que poder definir los aritméticos; (2) demostrar que los teoremas aritméticos son derivables de los principios lógicos mediante el único proceso válido, la deducción. Esto último obliga a especificar cuáles son los primeros principios lógicos y cuáles son las reglas de inferencia.

Y en vista de estos objetivos, Frege dará un primer paso: construir una lógica que le sea válida para su objetivo, una lógica del pensamiento puro, alejada de la influencia de la gramática y del lenguaje usual, para lo que debe crear un simbolismo adecuado. Y este primer paso, construcción de un simbolismo en el que poder expresar una lógica pura, independiente de la gramática, del lenguaje ordinario, de la psicología, es el que se contiene en el *Begriffsschrift* de 1879.

El paso siguiente se centra en la definición de número cardinal, que plasma en 1884 en *Fundamentos de la aritmética, un ensayo lógico-matemático sobre el concepto de número.* Es libro no menos fundamental de Frege, por la exposición, sin empleo de simbolismo ideográfico, de sus ideas y la crítica de las concepciones opuestas. Es libro del que Frege, quizá con un punto de amargura por el nulo éxito obtenido por *Bs*, señalará que para muchos no será otra cosa que un híbrido, ya que los filósofos lo estimarán como matemático y los ma-

BEGRIFFSSCHRIFT,

EINE DER ARITHMETISCHEN NACHGEBILDETE

FORMELSPRACHE

DES REINEN DENKENS.

VON

DR. GOTTLOB FREGE,

PRIVATDOCENTEN DER MATHEMATIK AN DER UNIVERSITÄT JENA.

HALLE A/S.

VERLAG VON LOUIS NEBERT.

1879.

GOTTLOB FREGE publicó, hace ahora cien años, *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (“Ideografía, un lenguaje de fórmulas para el pensamiento puro modelado en el lenguaje de la aritmética”), primer tratado sistemático de lógica matemática. No pretende ser un mero simbolismo del lenguaje ordinario, o un cálculo, sino una conceptografía que permita la traducción a signos que reflejen las relaciones entre los conceptos simbolizados mediante un manejo por reglas estrictamente especificadas. Es simbolismo apto para expresar el pensamiento puro.

temáticos como filosófico, y ni unos ni otros lo leerán.

Y por último, coronando su labor, los dos volúmenes, 1893 y 1903, de *Las leyes básicas de la aritmética*, donde afirma, en el primer tomo: "Con este libro llevo a ejecución un proyecto que tenía planteado desde mi *Begriffsschrift*, de 1879, y que empecé en mis *Fundamentos de la aritmética* de 1884".

Y, en medio, y al final, ensayos de precisión como "Función y concepto", "Sobre sentido y referencia", "Sobre concepto y objeto", "¿Qué es una función?", "Sobre la lógica en la matemática", "Investigaciones lógicas", en varias entregas... Ensayos que suelen reunirse bajo la etiqueta de *escritos filosóficos* o *lógicos-semánticos*.

Se señala en el prefacio de *Begriffsschrift* que existen dos tipos de juicios, los analíticos y los sintéticos. Frege estima que los aritméticos son juicios analíticos, contra el sentir kantiano, pero entiende por juicio analítico aquel que puede derivarse, en forma estrictamente lógica, de las definiciones. No se tiene en cuenta, aquí, el contenido de dicho juicio sino su derivabilidad. Es punto que Frege precisará en escritos posteriores, sin entrar en más detalles. Explica, a continuación, que la etapa inicial de su trabajo se centra en reducir el concepto de orden en una sucesión al de consecuencia lógica, para proceder desde allí al concepto de número. Para realizar esta tarea encuentra el lenguaje ordinario inadecuado. No sólo para esta labor. Agregará que una de las tareas de la filosofía debe consistir en liberar el espíritu humano de los errores que, en cuanto al concepto, presenta el lenguaje ordinario. En particular, debe eliminarse la confusa terminología entre "sujeto" y "predicado" en beneficio de "argumento" y "función". Para conseguir estos fines, dedica su atención a construir un lenguaje de fórmulas, a semejanza del aritmético, pero que permita un análisis lógico del razonamiento matemático, del pensamiento puro. Aplicado en particular a la aritmética en *Bs*, es posible generalizarlo tanto a la geometría como a la física; en general, se le muestra como herramienta útil para la filosofía en general. Y ello porque el lenguaje de fórmulas es al habla ordinaria como el ojo al microscopio. En este punto, Frege se remonta al intento de Leibniz de un *calculus ratiocinator*, reconociendo que Leibniz sólo se quedó en el intento, sin llevarlo a la práctica, como él ha logrado.

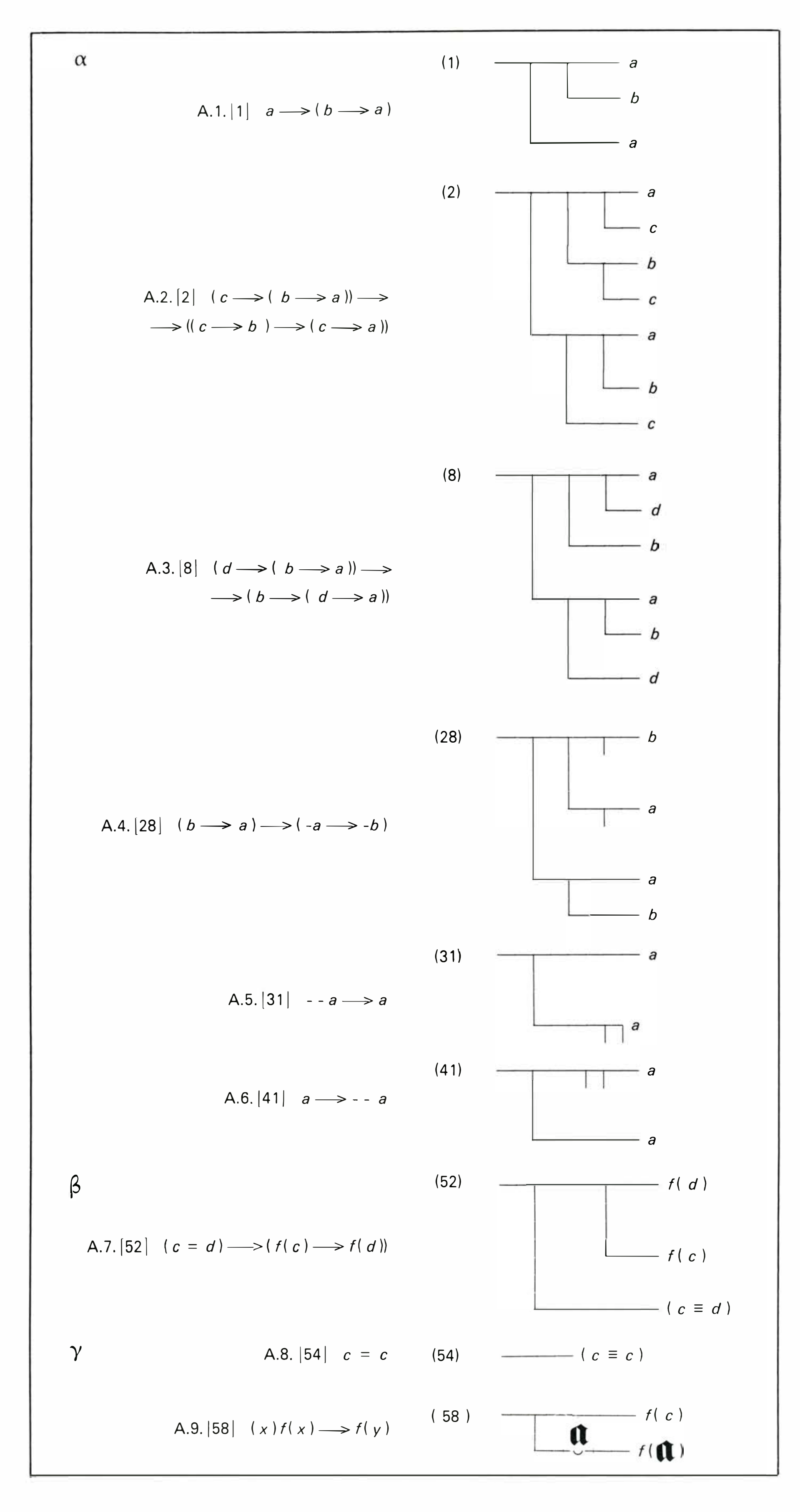
Estos objetivos conducen a escindir en dos grandes apartados su opúsculo: en el primero dará, primer capítulo, una descripción semántica de los símbolos que emplea; en el segundo capítulo, realizará una representación sistemática, deductiva, de algunos juicios del pensamiento puro. En otras palabras, expone, en el primer apartado, por vez primera, lo que hoy viene considerándose como lógica de primer orden –que incluye, por supuesto, la lógica proposicional–. Y ésta es la clave del tópico al estimar 1879 como el nacimiento de la lógica matemática. En el segundo apartado, o capítulo tercero, Frege aplicará su Ideografía para definir, por los medios estrictamente lógicos creados en la misma, la noción de "sucesión", y la de orden lineal o cadena, así como mostrar que el principio de inducción completa puede describirse por medio de su Ideografía. Deja para obra posterior la definición del soporte matemático de dicha sucesión, el número natural.

Debo precisar los motivos que Frege aduce para la creación de su Ideografía. No mera búsqueda de un simbolismo más o menos arbitrario y que refleje el lenguaje ordinario. Este tipo de simbolismo es, en el fondo, el propio del lenguaje matemático no formal, que por ello constituye una jerga especial formada por palabras y frases del lenguaje usual y por signos especiales. Jerga que no satisface a Frege por no ser válido para la búsqueda de precisión conceptual, aunque sea útil en la práctica. La Ideografía o conceptografía que logra debe permitir, por un lado, el que sea un cálculo lógico al estilo de lo preconizado por Leibniz, pero también que refleje el pensamiento puro y ello en el sentido de que, para Frege, el signo es inseparable pero consecuente al contenido que representa. Y este último punto deseo remarcarlo porque es convicción que se ligará siempre a Frege y a la que el matemático alemán permanecerá fiel: lo primero es el concepto; lo segundo, el signo con el cual se represente este concepto. Y es la creencia condicionadora de todo el trabajo de Frege lo que se ha venido a denominar su platonismo: el hombre no crea los conceptos, los aprehende; el hombre no crea sistemas matemáticos, sino que éstos preexisten conceptualmente al mismo; los contenidos conceptuales puros son independientes a que el hombre los perciba, los imagine, los piense... Un ejemplo que Frege expone en 1914 es que el pensamiento que se tiene en el teorema de Pitágoras es el mismo para todo ser humano, y su verdad es independiente de que sea o deje de ser pensado por algún individuo determinado. Y en lógica, en matemática, lo que importa es el pensamiento puro, no la génesis del mismo.

Es convicción que le lleva a oponerse a los métodos de Boole, radicalmente, porque en el fondo Boole parte en su labor de la construcción de un cálculo formal que permite ulteriores interpretaciones, distintas; para Frege ello equivale a partir del signo material para alcanzar el concepto. Y Frege insistirá en que tales cálculos, por su punto de partida, se mostrarán impotentes para la expresión, precisamente, de los conceptos y relaciones estrictamente lógicos. Es punto que explica cómo posteriormente Frege polemizará con Hilbert negándose a admitir el método axiomático formal.

Para Frege, así, lo primero es el contenido conceptual o de juicio; lo segundo, el signo con que pueden representarse tales contenidos o pensamientos. Y un contenido que no hace referencia, en momento alguno, a los aspectos psicológicos. De aquí su rechazo, incluso, de las concepciones de Husserl, porque la lógica hace referencia al pensamiento puro y no a función psíquica alguna. Como muy posteriormente precisará, una proposición lógica no es más que un signo compuesto con arreglo a una regla determinada; signo que posee un "sentido" que se mantendrá en cualquier lengua a la que se traduzca la proposición anterior. Y es este "sentido" el que Frege denomina *pensamiento*, independiente, por tanto, de la representación sensorial del mismo, de la actividad psicológica o espiritual más o menos subjetiva. Lógica frente a gramática, psicología, teoría del conocimiento.

Es por ello por lo que su Ideografía, aunque se inspire en la aritmética en cuanto al uso de letras como variables y como constantes, letras para expresar la generalidad y letras para representar aquello que posea un significado completamente determinado, no pretenda ser un cálculo sino una conceptografía que permita la traducción a signos que reflejen las relaciones entre los conceptos simbolizados mediante un manejo por reglas estrictamente especificadas. Y es punto que mantendrá frente a Schroder, quien en 1880 critica el empleo en el *Bs* de signos diferentes a los aritméticos para expresar el pensamiento puro, la lógica. Frege replicará en 1882 indicando, "En realidad, yo no he querido hacer un simple *calculus ratiocinator* sino una *lingua characterica (sic)* en el sentido de Leibniz". Y ello hasta el extremo de que si se partiera de un cálculo al estilo del álgebra lógica se está condenando a mantenerse en una especie de álgebra abstracta, vacía, mientras que puede concebirse una *lingua characterica* que no aboque en un cálculo por el mero cálculo. El cálculo no debe considerarse como otra cosa



PARA LA PRESENTACION DE SU SISTEMA, **Frege elige un total de nueve axiomas que, junto a las cuatro reglas, implican que dicho sistema de axiomas es completo, en el sentido de completitud de sistemas formales posterior (para la lógica de primer orden). La elección está presidida por el uso de cada una de las cuatro constantes lógicas primitivas. En α se dan los axiomas que Frege establece, con su numeración en *Begriffsschrift* (*izquierda*) simbolizados también con notación no fregeana (*derecha*). Junto a las reglas de separación o *Modus ponens* y la de sustitución, estos axiomas constituyen un sistema completo para el cálculo proposicional. En β se dan los axiomas de identidad y, en γ, el de cuantificación.**

que como un complemento de dicha *lingua*. Y es lo que Frege mostrará en *Bs.*

Tras indicar el manejo de letras, con su diferencia de constantes y variables, al modo de la aritmética, Frege pasa a determinar el elemento básico de su Ideografía: el juicio o la aserción. No el concepto, como venía siendo tradicional en lógico –salvo en la estoica–, o la clase como venía siéndolo en la línea booleana. Bien entendido que sin dar, en su Ideografía, una definición de lo que entender por juicio, ya que en la Ideografía se manejan fórmulas, es un lenguaje de fórmulas, por lo que sólo podrá establecer reglas o instrucciones para la manipulación de los signos que en ella se utilizan. Es advertencia válida para todos los restantes elementos lógicos que Frege introduce. Y es punto que no vio Russell, por ejemplo, en la crítica de las ideas de Frege que agrega como apéndice a *Los principios de la matemática*, 1903, donde señala que Frege no define qué sea el juicio, la negación. Sólo años después Frege indicará el concepto de proposición al que antes he aludido, pero apoyado ya no en el contenido de un juicio, sino en su distinción entre sentido y referencia.

Frente a los formalistas que llegan a identificar numeral y número, Frege distingue tres planos: expresión, contenido judicativo de esa expresión y aserción o juicio del contenido o pensamiento. Lo único que importa en la Ideografía es el contenido judicativo. "Los griegos vencieron a los persas en Platea" y "los persas fueron vencidos por los griegos en Platea" son dos expresiones diferentes, pero presentan el mismo pensamiento, el mismo contenido. Contenido que puede ser convertido en aserción, aunque sea independiente de tal aserción e incluso puedan existir contenidos que carezcan de la expresión asociada correspondiente. Ello conduce a rechazar la distinción entre sujeto y predicado, válida fundamentalmente para la expresión gramatical y no para el contenido judicativo ni para el conceptual. Es punto de partida que conduce a que la única diferencia que importa entre contenidos judicativos sea la que existe entre universales y particulares, porque dicha distinción lo es en cuanto a contenido conceptual y no sólo en cuanto a expresiones. De esta manera quedan fuera de la lógica las viejas distinciones entre juicios categóricos, hipotéticos, disjuntivos... Igualmente, conduce a admitir que la negación se aplica a contenidos de juicios y no a la sola expresión de los mismos, contenidos al que harán referencia, por modo exclusivo, las restantes constantes lógicas que explicitará Frege.

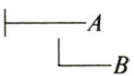
Desde este enfoque que diferencia radicalmente lógica de gramática y de teoría del conocimiento, Frege se ve obligado a rechazar la posibilidad de distinciones modales como tema propio de la lógica. Así, "es posible que la tierra choque algún día con otro cuerpo celeste" es una expresión en la cual quien la afirma no conoce las leyes de las cuales pueda seguirse la negación; en otras palabras, una distinción modal de posibilidades o de necesidad se refiere más al fundamento cognoscitivo que se tiene en el momento de enunciarla, que al contenido del juicio. Desde esta posición, aliada con la negativa a comenzar por el cálculo para alcanzar el concepto, se invalida cualquier construcción lógico-modal.

Para poder reflejar los tres planos: expresión– contenido judicativo –aserción de ese contenido, Frege crea un simbolismo especial. La expresión del contenido de juicio la representa por una mera abreviatura, una letra gótica que aquí reemplazo por letra latina como *A*. El contenido del juicio que se abrevia en la expresión por *A*, vendrá representado por un trazo horizontal o trazo de contenido: —*A*, que indica una simple conexión de ideas "sobre la cual el escritor no expresa si reconoce o no su verdad". Y la aserción de ese contenido, cuando puede convertirse en juicio, vendrá dada por un trazo vertical antepuesto a la línea de contenido: trazo de juicio. Así, el pensamiento "los polos magnéticos opuestos se atraen entre sí" es, como pensamiento, un contenido de juicio y puede representarse por '— *A*', donde *A* es la abreviatura, la expresión material del contenido: la conversión de ese pensamiento en juicio vendrá representado por '⊢— *A*'. De esta forma Frege indicará: "El signo ⊢— es el predicado común para todos los juicios".

Establecido el modo básico de representación de la aserción de un contenido judicativo, Frege pasa a describir las constantes lógicas primitivas. Serán, exclusivamente, las de generalidad, identidad de contenido y aquellas mediante las cuales se establezcan las relaciones entre contenidos judicativos elementales dadas por la condicionalidad y la negación. En otras palabras, los contenidos de pensamiento puro serán aquellos que pueden representarse con ayuda de la lógica proposicional y de la lógica de predicados con identidad. La descripción semántica de estas constantes lógicas primitivas y de las que pueden definirse en términos de las mismas, completará la primera parte, el primer capítulo, de este apartado.

La condicionalidad viene representada por un trazo de condición vertical que afecta al contenido de juicio y está estudiada de manera veritativo-funcional, en paralelo al condicional filónico, aunque probablemente Frege no conoció la construcción de Filón de Megara, por lo cual podría afirmarse que redescubrió independientemente lo que después se calificaría de "implicación material". Es punto, en cuanto a posibles influencias históricas, válido para el empleo de letras como variables ya utilizado por Aristóteles, en el que estimo que tales influencias no existieron en Frege, a pesar de sus llamadas a Leibniz.

Lo que expresa el esquema



es la exclusión del caso "*A* es negado y *B* afirmado", mientras que tiene lugar uno de los otros tres casos. Es decir, el esquema sólo excluye el caso en el que '*B*' es verdadero y '*A*' es falso, ya que: "Si *A* y *B* se ponen por contenidos que puedan hacerse juicios, hay las cuatro posibilidades:

1. Se afirma *A* y se afirma *B*
2. Se afirma *A* y se niega *B*
3. Se niega *A* y se afirma *B*
4. Se niega *A* y se niega *B*".

Frege precisa que la traducción, la lectura de esta representación del condicional por "si...", no es totalmente adecuada, salvo en el caso en que exista una relación causal –lo que es imposición excesivamente restrictiva– entre los contenidos representados por *A* y *B*. Lo cual viene a indicar que el condicional no tiene por qué darse precisamente con relaciones causales entre los contenidos de los juicios.

La escritura del condicional obliga a que el antecedente se escriba debajo y el consecuente arriba. La aparente arbitrariedad de esta colocación se justifica en el modo de inferencia fregeano. Las dos aserciones juntas, la dada por el condicional y la dada por el antecedente, aseguran la aserción '⊢—*A*'. Basta borrar el antecedente. Que la afirmación de las dos aserciones indicadas dé paso a la aserción del consecuente, constituyendo así un modo de inferencia, viene asegurado porque de los cuatro casos posibles antes reseñados, el tercero queda eliminado por el condicional, mientras que la aserción del antecedente elimina el segundo y el cuarto, por lo que sólo queda la primera posibilidad.

Esquema de inferencia que Frege propone como único afirmando que toda inferencia puede reducirse a este modo, al menos en la lógica proposicional: "empleo únicamente éste (modo de inferencia), al menos en todos aquellos casos en los que un nuevo juicio se deriva de más de uno" (parágrafo 6). Ello implica, sin embargo, un problema: la posibilidad y la conveniencia de que existan otros modos de inferencia. Un modo único quizá sea poco en el sentido de que las deducciones se hacen muy largas y exige un número mayor de premisas de las que partir, por lo que en ocasiones conviene introducir abreviaturas y nuevas reglas de inferencia; pero un único modo permite un rigor absoluto. Y este es uno de los objetivos fregeanos. Otro problema se centra en si Frege, de hecho, maneja con exclusividad este modo de inferencia que no es otro que el *modus ponens* o regla de derivación o separación. Y, de hecho, y a pesar de las afirmaciones constantes de que es su único modo de inferencia, se manejan otras tres reglas, aunque se regule su uso en cuanto al manejo de las letras; en particular, para el cálculo proposicional, una regla de sustitución para variables proposicionales.

Como segunda constante lógica primitiva para la lógica proposicional Frege caracteriza la negación con las palabras: "Si un pequeño trazo vertical se liga por debajo a la línea de contenido, se expresará la circunstancia de que *el contenido no tiene lugar*" (parágrafo 7). Así



significará "*A* no tiene lugar".

La combinación del condicional y de la negación permite definir los restantes conectivos proposicionales: así, la conjunción y la disjunción de la que Frege empleará únicamente su sentido no excluyente (frente a Boole, por ejemplo, quien manejaba la disjunción excluyente). Se reconoce, igualmente, el hecho de que el condicional podía tomarse como una constante lógica no primitiva, definido en términos, por ejemplo, de conjunción y negación. Sin embargo, la facilidad para el manejo de la inferencia es la que ha conducido a la elección de dicho condicional como constante lógica primitiva.

En la elección puramente convencional, por adecuada al simbolismo y a la inferencia, difiere Frege de la posición de Mc Coll, quien desde 1877 sostenía el punto de vista de que la lógica tenía como objetivo fundamental el cálculo proposicional en el que el principal conectivo debía ser algún tipo de implicación. Implicación convertida en elemento base de la lógica matemática a partir de Peano en el sentido de estimar que en la matemática todos los teoremas vienen establecidos bajo una estructura implicacional. Tesis de Peano retomada

1. $|1|\ a \longrightarrow (b \longrightarrow a)$ A.1.

2. $|2|\ (c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))$ A.2.

3. $(c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))) \longrightarrow ((b \longrightarrow a) \longrightarrow ((c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a)))$

   Por Sust. en A.1. $a \,/\, (c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a));\ b \,/\, b \longrightarrow a$

4. $|3|\ (b \longrightarrow a) \longrightarrow ((c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a)))$ M.P.2.3.

5. $((b \longrightarrow a) \longrightarrow ((c \longrightarrow (b \longrightarrow a)) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))) \longrightarrow (((b \longrightarrow a) \longrightarrow (c \longrightarrow (b \longrightarrow a)))$

   $\longrightarrow ((b \longrightarrow a) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))))$

   Sust. en A.2. $a \,/\, (c \longrightarrow b) \longrightarrow (c \longrightarrow a);\ b \,/\, c \longrightarrow (b \longrightarrow a);\ c \,/\, b \longrightarrow a$

6. $|4|\ ((b \longrightarrow a) \longrightarrow (c \longrightarrow (b \longrightarrow a))) \longrightarrow ((b \longrightarrow a) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a)))$ M.P.4.5.

7. $(b \longrightarrow a) \longrightarrow (c \longrightarrow (b \longrightarrow a))$

   Sust. en A.1. $a \,/\, b \longrightarrow a;\ b \,/\, c$

8. $|5|\ (b \longrightarrow a) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))$ M.P.6.7.

9. $(b \longrightarrow a) \longrightarrow ((d \longrightarrow b) \longrightarrow (d \longrightarrow a))$

   Sust. en 8 $c \,/\, d$

10. $((b \longrightarrow a) \longrightarrow ((d \longrightarrow b) \longrightarrow (d \longrightarrow a))) \longrightarrow ((c \longrightarrow (b \longrightarrow a)) \longrightarrow (c \longrightarrow ((d \longrightarrow b) \longrightarrow (d \longrightarrow a))))$

   Sust. en 8 $a \,/\, (d \longrightarrow b) \longrightarrow (d \longrightarrow a);\ b \,/\, b \longrightarrow a$

11. $|6|\ (c \longrightarrow (b \longrightarrow a)) \longrightarrow (c \longrightarrow ((d \longrightarrow b) \longrightarrow (d \longrightarrow a)))$ M.P.9.10.

12. $((b \longrightarrow a) \longrightarrow ((c \longrightarrow b) \longrightarrow (c \longrightarrow a))) \longrightarrow ((b \longrightarrow a) \longrightarrow ((d \longrightarrow (c \longrightarrow b)) \longrightarrow (d \longrightarrow (c \longrightarrow a))))$

   Sust. en 11. $a \,/\, c \longrightarrow a;\ b \,/\, c \longrightarrow b;\ c \,/\, b \longrightarrow a$

13. $|7|\ (b \longrightarrow a) \longrightarrow ((d \longrightarrow (c \longrightarrow b)) \longrightarrow (d \longrightarrow (c \longrightarrow a)))$ M.P.8.12

PRIMERAS SIETE FORMULAS que Frege demuestra en su Ideografía. Se las ha "traducido" a una notación más actual, señalando entre corchetes la equivalencia con las originales. Se indican las sustituciones simultáneas que deben hacerse en cada una de las líneas indicadas. Estas fórmulas sólo utilizan los dos primeros axiomas del cálculo proposicional y, además de la regla de sustitución, aunque Frege no la enuncia de modo explícito en parte alguna, se emplea la regla de derivación *Modus ponens* (que dice que, dada una fórmula condicional, probado el antecedente, se sigue el consiguiente).

por Russell al sostener que la matemática no es otra cosa que el conjunto de proposiciones de la forma "si *H* entonces *T*". Argumento que esbozará, nuevamente, para hacer aceptable axiomas como el de existencia del infinito, existencia no lógica, pero no asegurada sino puesta como antecedente de las proposiciones que constituyen la clase.

Descritos los elementos básicos para la lógica proposicional, se pasa a una introducción semántica de la teoría de la cuantificación (término que no es fregeano) con identidad. Es paso que va a dar origen a una continua revisión conceptual por parte de Frege y a una copiosa y amplia literatura en la filosofía de la lógica posterior. Por lo pronto, comienza con la introducción de un predicado de dos argumentos: la identidad. Introducción que pretende ser, y es, la primera definición de carácter estrictamente lógico de esta relación, dado que la definición aportada por Leibniz, por ejemplo, se apoyaba en el principio ontológico de los indiscernibles. Frege comienza su parágrafo 8 afirmando: "La identidad de contenido se diferencia de la condicionalidad y de la negación en que se aplica a nombres y no a contenidos".

No es, por tanto, identidad de objetos; de aquí que la definición fregeana se presente en metalenguaje. La expresión es

$$\vdash (A \equiv B)$$

que significa "el signo *A* y el signo *B* tienen el mismo contenido conceptual, de forma que en lugar de *A* se puede poner siempre *B* y viceversa". El juicio de identidad no es otra cosa que la aserción de que, en ocasiones, puede representarse el mismo contenido mediante dos maneras diferentes de determinación y Frege aclara este hecho con un ejemplo tomado de la geometría –de donde obtiene, realmente, este tipo de identidad– en el que uno y el mismo punto queda "determinado" de dos maneras distintas. Aunque no sea el ejemplo fregeano, puede estimarse que *A* se pone por "el centro de la circunferencia circunscrita a un triángulo rectángulo" y *B* por "el punto medio de la hipotenusa de un triángulo rectángulo". La identidad fregeana indica que las dos expresiones, diferentes, poseen el mismo sentido, el mismo contenido. Pero es juicio de identidad que no puede hacerse por la mera observación de dichas expresiones ni por la observación del contenido de cada una de ellas, sino que se obtiene después de realizar unas previas determinaciones o construcciones. El juicio,

la aserción que tras estas construcciones se logra, es que ambos puntos son el mismo. Y esta aserción es lo que expresa el juicio de identidad. Aserción que Frege califica como juicio sintético en terminología kantiana e incluso afirma que el mecanismo para alcanzar juicios de este tipo es un mecanismo válido para la creación de tales juicios.

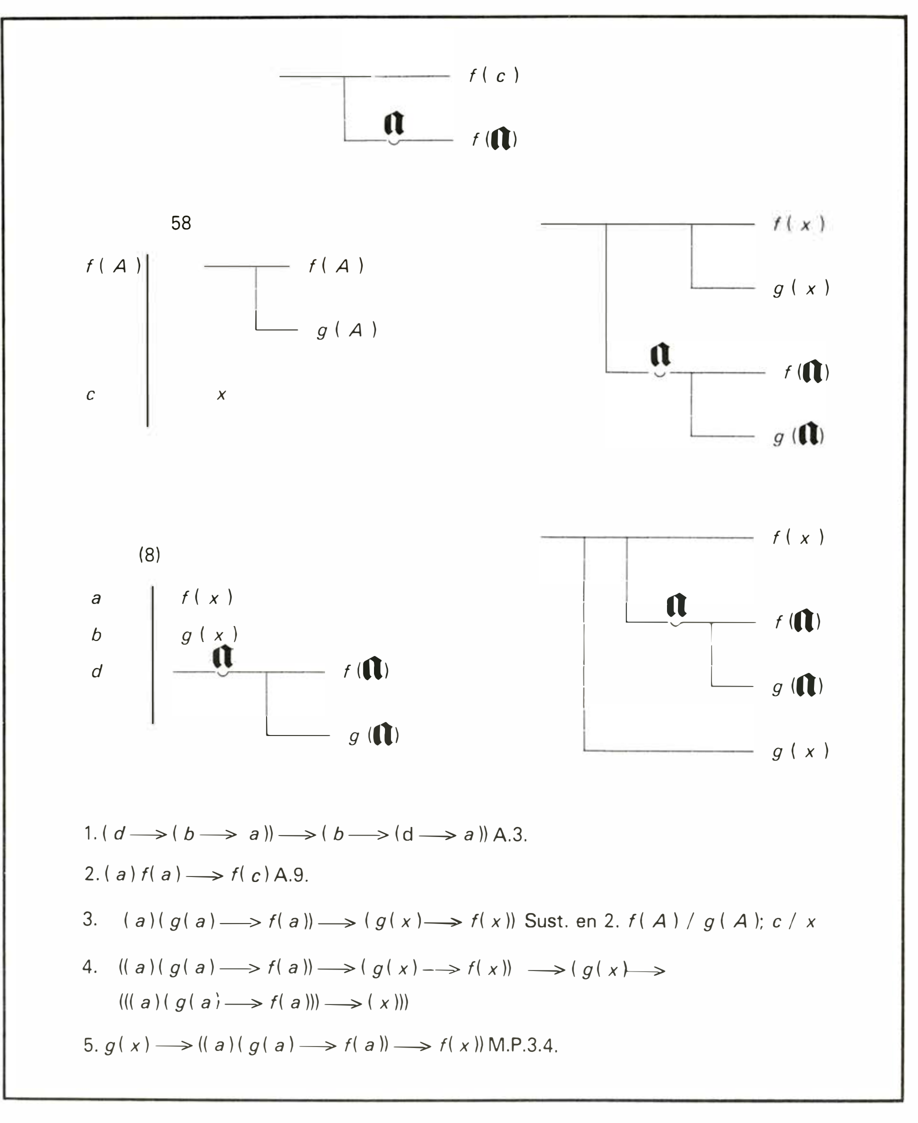
Lo que pretende hacer ver Frege es que el signo '≡' simboliza la identidad de los contenidos de las expresiones entre las cuales se intercala. Pero ello supone una dificultad, ya que si el juicio se refiere a los nombres, a las expresiones entre las cuales se intercala, entonces la formulación de dicha identidad no sería un problema lógico ni geométrico, sino estrictamente gramatical. Es dificultad vista por Frege posteriormene al *Bs*: ya que si en "$A \equiv B$" se hablara sólo de los signos 'A' y 'B' entonces "no interesarían ya las cosas mismas, sino tan sólo nuestra manera de referirlas". Es dificultad que pretenderá superar mediante el abandono del "contenido" que jamás ha definido, creando la distinción entre "sentido" y "referencia". Distinción que le conduce a la posterior supresión de este tipo de identidad en beneficio de la igualdad matemática. El problema de la identidad en lógica vuelve a renacer como problema, aunque en general se vuelva a la concepción de identidad no entre nombres sino entre objetos. Por otro lado, la distinción fregeana entre "sentido" y "referencia" va a constituir el punto de apoyo de gran parte de la semántica filosófica de este siglo, aunque es distinción ausente del *Bs*.

Para poder establecer la nueva constante lógica de generalidad, la cuantificación, Frege pasa a elaborar lo que considera uno de sus hallazgos fundamentales: el de función. Concepto que adopta del hacer matemático, así como la notación empleada, pero que estima de un carácter más general, al igual que adoptara de la aritmética la distinción de letras para variables y letras para constantes; y adopción, pero ampliada, porque estima que ambos casos poseen un carácter más general que el que les sirve de modelo (parágrafo 10).

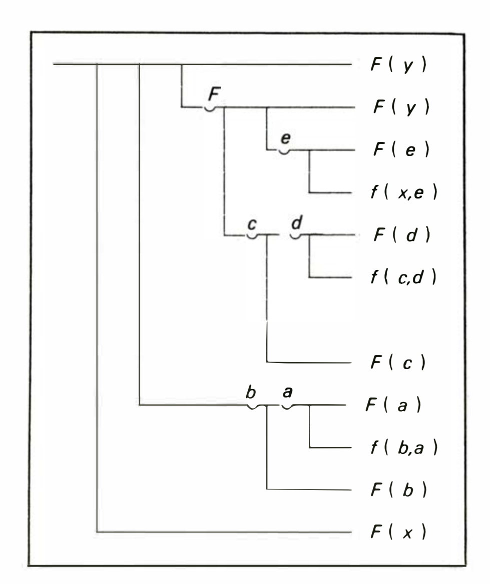
Y para hacer más claro, o aceptable, esta introducción, Frege comienza por unos ejemplos químicos con los cuales muestra cómo analiza una proposición no en sujeto y predicado sino en argumento y función. Sea una expresión como "Juancito come hierba" (el ejemplo no es de Frege, claramente). Si en lugar de "Juancito" se pone "Jaimito", la expresión seguirá siendo válida. Se puede reemplazar el término "Juancito" por otros términos o, con generalidad, por un lugar vacío: "(   ) come hierba", y ello de manera tal que, al cubrir ese espacio vacío por un término conveniente se tenga la expresión completa que podrá o no ser judicable. Y lo será cuando el término sea conveniente, en cuyo caso dicho término poseerá la propiedad indicada por la otra parte de la expresión; en este ejemplo, "Juancito" poseerá la propiedad de comer hierba. Todos aquellos términos que permitan cubrir el espacio vacío constituirán los *argumentos*, mientras que la propiedad que los mismos poseen, la de "comer hierba", constituye la *función* para tales argumentos. Si ahora se toma la expresión "Juancito ama a Juancita", en lugar de "Juancito" y "Juancita" pueden colocarse otros términos por argumentos, por lo que la expresión general tendría dos espacios vacíos "(   ) ama a (   )" y la función "ama a" será una función de dos argumentos; como "(   ) es menor que (   )". El proceso puede continuar generalizándose para obtener funciones pluriargumentales.

Los espacios vacíos se representarán por letras entre paréntesis, como indeterminadas, mientras que la propia función se representará, igualmente, por una letra. Representación que Frege hace por '$\emptyset$ (A)' para la función de un argumento y '$\emptyset$ (A, B)' para la de dos argumentos. Si al reemplazar "convenientemente" la letra entre paréntesis resulta que el contenido obtenido es capaz de ser convertido en juicio, en aserción, entonces es que el argumento satisface la función, es decir, posee la propiedad determinada por la misma. Y ello es lo que expresa la aserción correspondiente que vendrá representada por '$\vdash$ $\emptyset$ (A)' que indica "A tiene la propiedad $\emptyset$"; mientras que la aserción '$\vdash$ $\emptyset$ (A, B)' puede leerse como "B está en la relación $\emptyset$ con A". Si se intercambian los argumentos A y B entre sí en una función de dos argumentos, puede no tenerse una aserción y, de tenerse, puede no coincidir con la original.

En la Ideografía, Frege no da definicio-



FORMA EN QUE FREGE demuestra el modo de inferencia Bárbara cuando la premisa menor, $g(x)$, posee un contenido particular (*arriba*). En la parte inferior se indica, en símbolos actuales, ese mismo modo de inferencia indicando los axiomas que se utilizan y donde realizar las sustituciones debidas.



PRINCIPIO DE INDUCCION COMPLETA, que, en Frege, adopta la siguiente formulación: "Si *x* tiene la propiedad *F* que es hereditaria en la *f*-suçesión, e *y* sigue a *x* en la *f*-sucesión, entonces *y* tiene la propiedad *F*". La ilustración reproduce, sin utilizar ninguna abreviatura, la forma que adquiere el principio en la Ideografía. Indica, a la vez, las dificultades del simbolismo fregeano.

nes, sino meramente reglas de manejo de fórmulas. De aquí que Frege no establezca definición alguna de lo que sea una función. Tema al que volverá posteriormente en varios ensayos como los que he mencionado antes. Tampoco indica cómo establecer la "conveniencia" en el reemplazo de una indeterminada en el espacio vacío para obtener la aserción de una función. Igualmente, se mantiene alejado de los conceptos matemáticos de dominio y contradominio o recorrido que podrían entrañar un aspecto extensional al interpretarse como conjuntos o clases, aunque ambos estén prácticamente explicitados, dado que aquellos indeterminados que reemplacen en lugar vacío constituirán el dominio de la función, mientras que los que sirvan para dar la aserción constituirán el recorrido de la misma. Será posteriormente cuando introduzca la noción de "recorrido de valores", obligado para la caracterización del número natural como cardinal. Y así, en el parágrafo 9 de *Las leyes básicas de la aritmética,* de 1893, afirmará: "La introducción de la forma de simbolización de extensión de conceptos (recorrido de valores) me parece que es una de las más fructíferas extensiones de mi ideografía que he hecho desde mi primera publicación sobre esta materia". Introducción que, sin embargo, le alejará, como reconoce en 1910, del carácter estrictamente formal explicitado en *Bs.* Alejamiento en el que ve una posible causa de la aparición de las antinomias, ligadas, precisamente, a la noción de "extensión de concepto".

Además, Frege, en la explicación que hace de la noción de función viene a indicar la posibilidad de la cuantificación de predicado. Sus palabras: "también podemos considerar a ∅ (*A*) como una función del argumento ∅" (parágrafo 10). Una exposición sistemática de la lógica como cálculo puede no hacer uso de este tipo de cuantificación. Frege la emplea en varias ocasiones, fundamentalmente en su último capítulo, al tratar de definir, por medios estrictamente lógicos, la noción de sucesión y el principio de inducción completa.

Es el análisis de una proposición en letra funcional y argumento el que permite superar a Frege la clásica distinción, de origen gramatical, entre sujeto y predicado. Análisis por el cual puede establecer uno de los logros más definitivos de la lógica matemática: la teoría de la cuantificación. Siguiendo con la función, puede ocurrir que todo término que se reemplace en el argumento de una función posea esta propiedad; y la expresión de este hecho viene simbolizada por Frege dotando al trazo de contenido de una concavidad en la cual se coloque una letra gótica minúscula, la misma que debe situarse en el argumento de la función. Es la aserción de que *todo* cumple la propiedad ∅, y viene representada por

$$\vdash\!\overset{\mathfrak{a}}{\smile}\!\!-\; \emptyset(\mathfrak{a})$$

Y según se vayan reemplazando las letras góticas pueden irse obteniendo aserciones particulares. La concavidad con la negación permite la expresión de aserciones existenciales, en el sentido de la lógica clásica.

El manejo del cuantificador exige algunas condiciones. Por lo pronto, debe estar sometido a que cualquier sustitución que pueda hacerse en una función tiene que dar un contenido que pueda convertirse en juicio: "si una combinación de signos que siguen a un trazo de contenido puede convertirse en juicio, entonces esa posibilidad permanece inalterada por una sustitución" (parágrafo 11). Además, la letra gótica situada en la concavidad del trazo de contenido aparece como una variable ligada y, por ello, es diferente a una variable libre (términos que no son de Frege, sí su concepto). La concavidad "delimita el alcance que cubre la generalidad indicada por la letra. La letra gótica retiene un significado fijo sólo dentro de su alcance propio; dentro de un juicio, la misma letra gótica puede ocurrir en alcances diferentes, sin que el significado atribuido a ella en un alcance se extienda a ningún otro". Igualmente, el alcance de una letra gótica puede incluir otras, pero en este caso debe tomarse la precaución de que las letras deban ser elegidas diferentes, sin confundirlas.

Frege reconoce que hay casos en los cuales la sustitución no es posible, como cuando el contenido del juicio está bajo el alcance de la letra gótica, es decir, cuando esta letra tiene por alcance a la totalidad de la expresión en la cual interviene. En este caso, sobra expresar la cuantificación. Y con deseo de distinguir estas dos posibilidades introduce otra notación para expresar la generalidad: emplear una letra itálica no precedida de cuantificación; así, '$\vdash X(a)$'. Pero, condición que Frege impone, debe quedar abierta la posibilidad del paso de una a otra expresión.

"Por ejemplo, en lugar de

$$\vdash\!\!-\; X(a)$$

podemos escribir

$$\vdash\!\overset{\mathfrak{a}}{\smile}\!\!-\; X(\mathfrak{a})$$

si *a* ocurre únicamente en los lugares de argumento de *X* (*a*)".

No voy a detenerme más en lo que no es otra cosa que una descripción de las reglas básicas de la lógica cuantificacional o de predicados. Reglas con las que cualquier tratado comienza hoy día. Me bastaba indicar que ellas están expuestas con nitidez y rigor en este libro de 1879, cuando todavía en 1894 Peano estimaba que la teoría de la cuantificación era complicada, por lo que se limitaba a dar, no la teoría, sino ejemplos de la misma, aunque posteriormente, en 1897, la reestableciera e incluso diera nombres a los distintos tipos de variables bajo los términos de real y aparente o libres y ligadas. Cuantificación que fue redescubierta no sólo por Peano, sino por Mitchell y por Peirce independientemente de Frege y entre sí.

El hecho es que Frege agrega a las dos reglas del cálculo proposicional (la explícita del *modus ponens* y la implícita de sustitución simultánea de variables proposicionales) otras dos para el cálculo de predicados. La tercera regla no es otra que la señalada más arriba: el paso de la expresión de la generalidad sin cuantificador a la expresión con cuantificador; es la que hoy se denomina "regla de generalización". La cuarta regla la enuncia Frege inmediatamente tras la anterior:

"Es claro también que de

$$\begin{array}{l}\vdash\!\!\!\top\!\!-\; \emptyset(a)\\ \quad\llcorner\!\!-\; A\end{array}$$

podemos derivar

$$\vdash\!\!\overset{\mathfrak{a}}{\smile}\!\!\!\begin{array}{l}\phi(\mathfrak{a})\\ A\end{array}$$

si *A* es una expresión en la que *a* no ocurre y si *a* ocupa únicamente posiciones en los lugares de argumento de ϕ (*a*)".

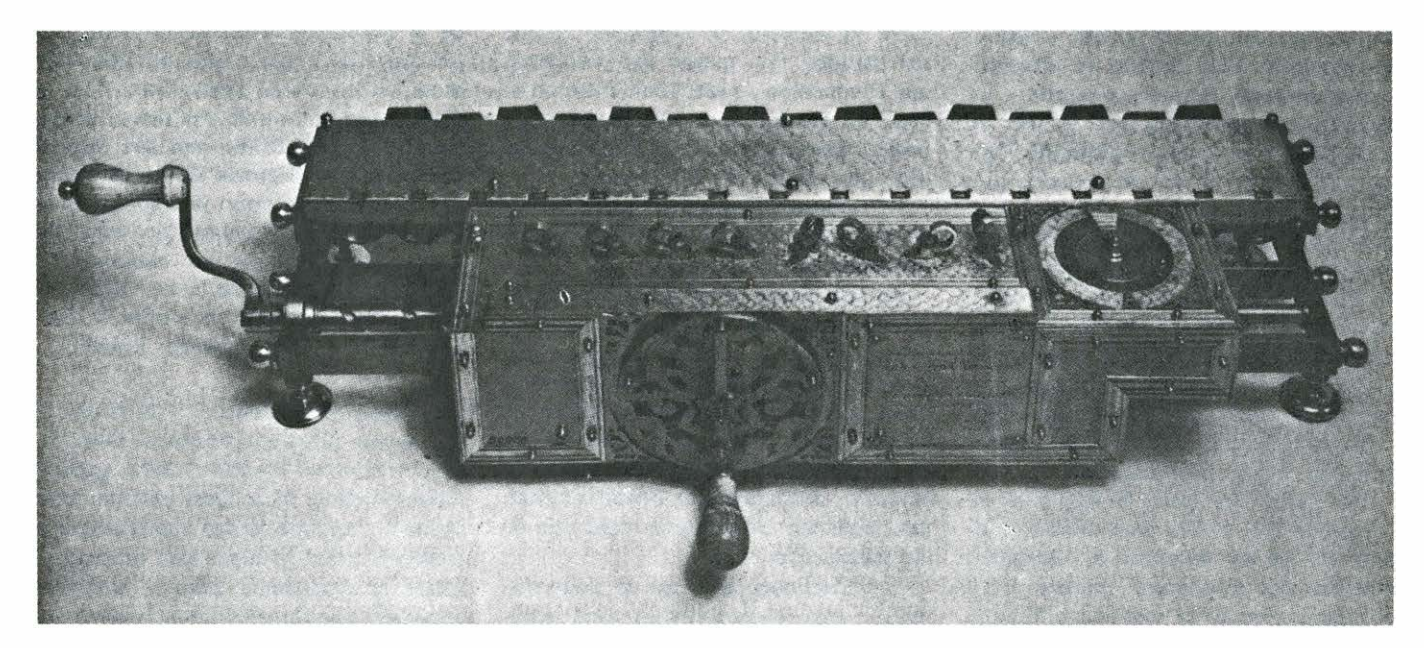
Podría indicarse que las cuatro reglas no están establecidas como modos de inferencia salvo el *modus ponens*, porque las mismas se justifican por modo exclusivo en cuanto al uso simbólico. Y ello quizás aparezca más claro en el capítulo 2, en la exposición sistemática de la lógica, ahora como cálculo, presentación que no es la de una teoría formal axiomática, sino más bien presentación sistemática de la lógica con ayuda de la Ideografía, de un simbolismo adecuado.

Realizada la descripción del simbolismo y, con él, una serie de principios del pensamiento puro, Frege pasa a dar un tratamiento sistemático en su capítulo dos. Es la exposición del cálculo proposicional y del cálculo de predicados con identidad que, según acabo de indicar, aparece como exposición sistemática y no como una mera construcción formal axiomática, como mero cálculo. Sin embargo, tal construcción puede enfocarse de esta última manera tras la exposición de la morfología correspondiente y entonces queda englobada como una exposición más formal axiomática. Es el enfoque muy posterior a Frege, pero que ha prevalecido en la exposición de las teorías formales, especialmente desde los entornos de 1930, tras los teoremas de incompletitud de sistemas formales; concepto de sistema formal inexistente en Frege e incluso opuesto a su concepción de la lógica como he venido indicando.

El objetivo de Frege es: "Parece natural derivar los juicios más complejos de los más simples, no para hacerlos más ciertos, lo que sería innecesario en muchos casos, sino en función de hacer manifiestas las relaciones de unos juicios con otros. Conocer meramente las leyes no es evidentemente lo mismo que conocerlas junto con las conexiones que tienen con otras" (parágrafo 13). De aquí que pueda llegarse a un pequeño número de leyes que, junto a las contenidas en las reglas, posean el contenido de todas las demás, aunque en estado no explicitado, no desarrollado. Y Frege señala cómo, para hacer ese desarrollo y esas conexiones visibles, es el modo deductivo de presentación el más adecuado. Bien entendido que no para hacer valer la certeza del pensamiento puro, o para construir juicios u obtener conceptos a partir de otros conceptos componiéndolos entre sí, sino para manifestar las relaciones de los juicios de ese pensamiento puro entre sí. El método deductivo como meramente organizativo, no como método de construcción de sistemas ni de obtención de nuevas proposiciones. Es una concepción, nuevamente, encontrada respecto al enfoque estrictamente formal axiomático del hacer matemático.

El problema se centrará, entonces, en la elección de aquellas leyes que se adoptan como punto de partida; leyes que posean en sí todo el contenido de las restantes, junto a las reglas, y que por ello mismo permitan obtenerlas. Frege reconoce que pueden adoptarse distintos conjuntos de leyes de partida. Pero con esta breve presentación, perfila con relativa claridad lo que constituye un sistema lógico deductivo a la vez que establece, esta vez con radical precisión, la distinción entre ley y regla.

En este punto Frege se aparta mucho más de la tendencia booleana, por ejemplo. Como indicará posteriormente, las expresiones "por lo tanto", "por consiguiente"... indican inferencias, pero nunca explicitan las leyes por las cuales se realizan dichas inferencias. Y lo que importa en el pensamiento puro no es sólo asegurarse de la certeza de la proposición inferida, sino poner de relieve la justificación de dicha inferencia. De lo contrario, incluso pueden emplearse palabras como las anteriores sin que, de hecho, se tenga inferencia alguna. De aquí que en el *Bs* se expongan unos cuantos pensamientos puros en forma de cálculo, pero no en el aspecto booleano, no como mero algoritmo manipulador al estilo del álgebra lógica o de la suma y el producto, sino al estilo de un algoritmo como conjunto de reglas que regulen el paso de una o dos proposiciones a una tercera. Con ello lo que se asegura es la exactitud del proceso demostrativo y no la certeza de la proposición obtenida. De aquí su permanente afirmación de que sólo hace uso de un único modo de inferencia, como regla, punto que fue descuidado, por ejemplo, por los constructores del álgebra lógica



MAQUINA CALCULADORA diseñada por Gottfried Wilhelm Leibniz, para multiplicar y dividir mediante la adición y la sustracción iterada. Quizá no llegó a construirse en su tiempo. El modelo de arriba es una versión aplicada siguiendo sus indicaciones, construido en 1923. Frege pretende que su Ideografía sea un cálculo lógico, pero también que refleje el pensamiento puro y ello en el sentido de que, para él, el signo es inseparable pero consecuente al contenido que representa. Los elementos fundamentales de la máquina eran ocho cilindros que portaban dientes de distinta longitud.

ALGEBRA DE BOOLE	CALCULO DE PROPOSICIONES
U (CONJUNTO UNIVERSO)	V (VERDADERO)
Ø (CONJUNTO VACIO)	F (FALSO)
a,b,c,... (CONJUNTOS, SUBCONJUNTOS, ELEMENTOS)	*p,q,r,...* (PROPOSICIONES)
$a \cup b$ (REUNION: TODO *a* Y TODO *b*)	$p \vee q$ (DISYUNCION: O BIEN *p* SOLO O *q* SOLO, O AMBOS, SON VERDADEROS)
$a \cap b$ (INTERSECCION: LO QUE *a* Y *b* TIENEN EN COMUN)	$p \bullet q$ (CONJUNCION: AMBOS, *p* Y *q*, SON VERDADEROS)
$a = b$ (IDENTIDAD: *a* Y *b* SON EL MISMO CONJUNTO)	$p \equiv q$ (EQUIVALENCIA, SI Y SOLO SI *p* ES VERDADERO, *q* ES VERDADERO)
a' (COMPLEMENTARIO: RESTO DEL CONJUNTO UNIVERSO QUE NO ES *a*)	$\sim p$ (NEGACION: *p* ES FALSO)
$a \in b$ (INCLUSION: *a* ES ELEMENTO DE *b*)	$p \supset q$ (IMPLICACION: SI *p* ES VERDADERO, *q* ES VERDADERO)

GEORGE BOOLE, desde 1847, y partiendo del hacer matemático, había realizado la construcción de una álgebra lógica, esto es, aplicar el álgebra para fundamentar la lógica. Boole parte de las nociones de clase, elemento de clase y operaciones con clases. El enfoque estriba en que las leyes del pensamiento, las leyes de la lógica, deben ser del mismo tipo que las que gobiernan el álgebra; es decir, la validez de los procesos del álgebra no depende de la interpretación de los signos, sino de las leyes de combinación de los mismos. Con estas ideas, Boole algebriza la lógica obteniendo un sistema algebraico que es, en términos actuales, un retículo booleano. Dicho retículo, como álgebra lógica o álgebra simbólica, presenta dos interpretaciones: por un lado, un álgebra de clases; por otro, un álgebra proposicional.

(y no sólo por Euclides, quien no explicita las reglas de derivación de las que hace uso).

Para la presentación de su sistema Frege elige un total de nueve axiomas que, junto a las cuatro reglas ya mencionadas, implican que el sistema fregeano es completo —en el sentido de completitud de sistemas formales posterior—. La elección está presidida por el uso de cada una de las cuatro constantes lógicas primitivas. Y de tal manera que Frege no las explicita en bloque, sino que va derivando de cada grupo de axiomas algunas de sus consecuencias. Así, las fórmulas que enumera por (1), (2) y (8) constituyen los axiomas de condicionalidad; a ésta se agrega la negación y los axiomas que regulan ambos conectivos vienen dados por las fórmulas (28), (31) y (41). Son los seis axiomas que regulan el cálculo proposicional.

Dos leyes más para la identidad de contenido (52) y (54). Esta última no es otra que la ley reflexiva de identidad, mientras que la primera expresa la ley de sustitución de la identidad. Es una expresión que manifiesta el carácter extensional de la ley de identidad en el sentido de que se pueden reemplazar dos variables de individuo entre sí en cualquier contexto sin variar el valor veritativo de las expresiones en que se intercambian. Axioma que puede interpretarse extensionalmente en el sentido de que todo lo que se predique de un símbolo, de un argumento, tiene que predicarse del símbolo, del argumento al que sea idéntico. Ya apunté los problemas que el concepto de identidad entre nombres, no entre objetos, implicaba y los posteriores cambios que el mismo Frege tuvo que realizar en este campo.

Por último, Frege agrega una fórmula para regular la cuantificación, y que constituye su fórmula (58).

De esta forma, Frege se permite escindir su sistema en varias zonas que van superponiéndose: cálculo implicacional y afirmativo; cálculo proposicional completo; cálculo de predicados; lógica de primer orden. Incluso al final de *Bs*, como apéndice, incorpora un cuadro en el que va señalando qué fórmulas se han ido utilizando para obtener cada una de las establecidas en *Bs*.

La independencia o no de unas respecto a otras leyes tomadas como axiomas no parece plantearse ni siquiera como problema. Menos aún, los problemas de completitud o consistencia. Ello exigiría un cambio epistemológico y admitir como punto de partida no el concepto, sino el sistema formal. Y ya he indicado varias veces que ello supondría una posición opuesta a la sostenida por Frege. En cuanto a la independencia, como demostró Lukasiewicz en 1930, la tercera ley de condicionalidad es consecuencia de las dos primeras y el sistema proposicional puede reducirse a sólo tres axiomas, siendo los dos primeros los de Frege y reemplazando los otros tres por el axioma

$$\vdash (a \rightarrow b) \rightarrow (-b \rightarrow -a),$$

escrito en notación, por supuesto, no fregeana.

Construida la Ideografía, Frege se cree en condiciones de pasar a su objetivo central: fundamentar la aritmética. Y, para ello, en *Bs*, en su tercer y último capítulo, muestra la posibilidad de este lenguaje de fórmulas para el pensamiento puro, para la construcción de la aritmética con independencia a cualquier contenido de los sentidos o de cualquier intuición *a priori*. Especialmente, se limita a una de las claves en dicha aritmética: la noción de sucesión que entraña una teoría de ordinales finitos. En particular introduce las nociones de antecesor, antecesor propio, correspondencia unívoca y orden lineal (fórmulas, respectivamente, 76, 99, 115 y 133, con la que termina *Bs*). Y trata de mostrar que la inducción completa puede ser incluida, con estas nociones, bajo el proceso de deducción, bajo el único modo de inferencia deductivo válido.

No voy a tratar, aquí, de este aspecto final de *Bs*, que sólo constituye un inicio de la fundamentación de la aritmética por parte de Frege y que, por ello, éste modifica en obras posteriores. Voy a mencionar, por modo exclusivo, no ya algunas proposiciones, con sus demostraciones, sino algunas de las definiciones que indiquen tanto la potencia de la Ideografía como el peligro de dicha potencia, especialmente las nociones de "propiedad hereditaria" y "antecesor de", por su influencia en la obra de Carnap, Quine y, por supuesto, Russell.

La noción de antecesor la trata de caracterizar Frege en términos, siempre, de una relación binaria *f*, dada, aunque indeterminada. Para hacer claro su pensamiento, aunque estima que el lenguaje ordinario, aquí, muestra aún más claramente su impotencia para la auténtica expresión de estas nociones, Frege pone un ejemplo. Sea $f(x, y)$ la relación que indica que *y* es hijo de *x*, y sea *Fa* la circunstancia, la propiedad de que *a* es un ser humano. Entonces cabe indicar

que la propiedad F es hereditaria respecto a la relación f, ya que si y es hijo de x, y x es un ser humano, entonces y también es un ser humano. La propiedad de "ser humano" es hereditaria respecto a la relación "ser hijo de", ya que el hijo de un ser humano es un ser humano. Es la clave de la afirmación de que una propiedad F sea hereditaria respecto a una f-sucesión. Que en formalización no fregeana sería:

$$Her\ F = (b)(Fb \rightarrow (a)(f(b, a) \rightarrow Fa))$$

Inmediatamente pasa a caracterizar la relación "x es antecesor de y en la f-sucesión" o "y sigue a x en la f-sucesión", en términos de la noción de propiedad hereditaria F. Si se tiene la relación $f(x, y)$ y si se verifica que cualquier elemento que esté f-relacionado con x posee la propiedad F, entonces y también posee la propiedad F, cabe decir que y es sucesor de x. Es lo que constituye la Def. 76 en Frege y que voy a representar en la forma

$$xAy = (F)(HerF \rightarrow [(a)(f(x, a) \rightarrow Fa) \rightarrow Fy)]$$

Ello no significa que la f-sucesión comience por un x determinado. Es lo que precisará Frege en su Def. 99, donde establece: "Si z es idéntico con x o sigue a x en la f-sucesión, entonces digo: 'z pertenece a la f-sucesión que comienza con x' o 'x pertenece a la f -sucesión que termina con z' ".

Lo que quiero destacar es que en la definición de antecesor, tiene que utilizar la cuantificación de predicado. Interviene ya el "para *toda* propiedad F". Le es inevitable para poder caracterizar la inducción completa que enuncia en los términos: "Si x tiene la propiedad F que es hereditaria en la f-sucesión, y si y sigue a x en la f-sucesión, entonces y tiene la propiedad F ".

$$Fx \rightarrow [(HerF \rightarrow (\ xAy \rightarrow Fy)]$$

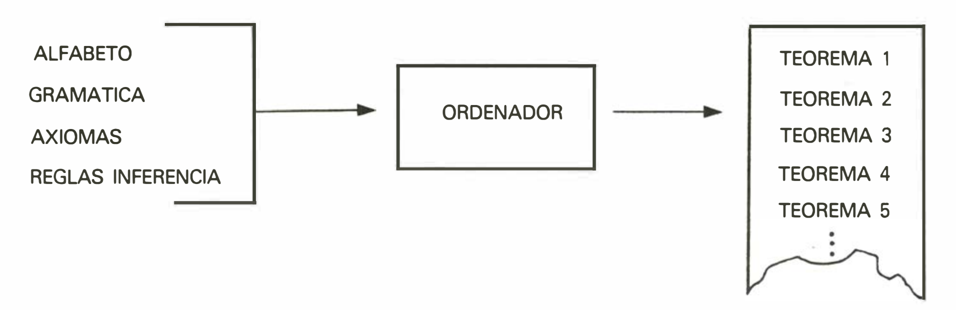
Como el punto de partida, la aritmética, es el mismo, Frege ha de coincidir aquí, anticipándose, con Dedekind y Peano. La diferencia estriba en que estos dos últimos tratan de escribir las propiedades aritméticas fundamentales, la estructura que subyace a las mismas, limitándose a establecerlas de modo axiomático, encubierto en Dedekind, explícito en Peano. Frege quiere ir más allá y tratará de definir el soporte del razonamiento inductivo matemático, el número natural, ya que el concepto de sucesión introducido en *Bs* se le muestra como más general que el de cadena (de hecho, el ejemplo aducido permite la ordenación no lineal, sino arbórea), siendo esta última propiedad la caracterizadora de los naturales. Para ello debe introducir la extensión del concepto, adoptando como paradigma la definición por abstracción mediante la relación de equivalencia. Y en esta línea coincide también con Cantor, quien la inicia hacia 1874 en correspondencia con Dedekind. Este ir más allá respecto a Dedekind y Peano lo manifestará Frege en 1884, en los *Fundamentos de la aritmética*, ya que el *Bs* no contiene las nociones de clase o de extensión de conceptos, manteniéndose en un plano de carácter estrictamente formal como comentará Frege en 1910. Plano formal en el que cree ver la posibilidad de evitar antinomias.

Arrepentimiento tardío, pero que tampoco hubiera impedido dicha aparición, porque la misma se encuentra larvada en *Bs*, en la afirmación que he citado de que una función $\emptyset\ (A)$ puede servir como argumento para $\emptyset$, junto al uso sin limitaciones del cuantificador funcional. Es lo que aparece en la definición de antecesor, en algunas demostraciones de propiedades de las sucesiones, donde roza la antinomia, como en la fórmula (91), que explicita:

"De la proposición (a), 'todo resultado de una aplicación del proceso f a x tiene la propiedad F', puede inferirse, para toda F, que todo resultado de una aplicación del proceso f a x tiene la propiedad F. De aquí también se puede inferir de la proposición (a) y la proposición de que la propiedad F es hereditaria en la f-sucesión, para todo F, que todo resultado de una aplicación del proceso f a x tiene la propiedad F."

Origen larvado de las antinomias, de la impredicatividad o círculo vicioso, es el "*todas* las propiedades" junto a la admisión de que una propiedad pueda ser propiedad de ella misma. En otras palabras, se centra en la potencia de la Ideografía para poder expresar cualquier tipo de cuantificación, sin limitación alguna y no, como por otro lado es correcto, en la afirmación de encontrarse en sólo la extensión de un concepto. Potencia que sólo posteriormente precisará Frege al distinguir entre "función" y "concepto" de forma tal que impida que un predicado pueda ser predicado de sí mismo, al admitir como argumentos sólo objetos.

He reiterado que el *Begriffsschrift* constituye el primer tratado sistemático de lógica matemática. Al paso, que otros autores (fundamentalmente los seguidores de la línea algebrizadora de la lógica, opuesta a la fregeana, iniciada por Boole y seguida por Peano y su escuela) fueron avanzando en líneas parecidas, redescubriendo y reelaborando elementos de esta lógica, como por ejemplo la clave de la misma, la cuantificación. Elementos que plasmará Russell en la obra de propósitos de 1903, *Los principios de la matemática* y culminarán Russell y Whitehead en *Principia Mathematica*. Pero, como primer trabajo que condensa en sí todas las innovaciones de la lógica matemática, el *Bs* constituye un libro modelo. En él, y en los dos primeros capítulos, se exponen con total precisión y nitidez como he intentado indicar, no sólo la cuantificación y las reglas de su uso, sino la previa distinción del concepto de función lógica de uno o varios argumentos, la distinción y uso sistemático de letras para variables y constantes, qué sea un sistema lógico deductivo y una exposición de dicho sistema donde las derivaciones se realizan atendiendo exclusivamente a la forma de las expresiones, se distingue entre ley y regla (distinción que sólo desde 1931 puede volver a suprimirse gracias a los teoremas de completitud y de Herbrand, aunque a costa



SISTEMAS FORMALES ideados por David Hilbert. Contienen un algoritmo que verifica de un modo mecánico la validez de todas las pruebas que puedan construirse a partir de un sistema. El sistema formal consta de: un alfabeto de símbolos con ayuda de los cuales pueda escribirse todas las proposiciones; una gramática que determina cuál es la forma como deben combinarse los símbolos; un conjunto de axiomas, o principios adoptados sin demostrar, y reglas de inferencia para reducir los teoremas a partir de los axiomas. Se obtienen los teoremas al escribir todas las proposiciones gramaticales posibles en el sistema y verificarlas para determinar cuáles son los concordes con las reglas de inferencia y, por tanto, válidas. Por ser esta operación realizable a través de un algoritmo puede llevarla a cabo un ordenador. En 1931 Kurt Gödel demostró que todos los sistemas formales venían a ser incompletos.

de sacrificar el concepto estricto de sistema lógico deductivo). Desde este punto de vista, como libro de centenario de la lógica matemática, y a pesar de que contiene muchos otros puntos de indudable valor, el tópico puede admitirse respecto al "nacimiento" de la lógica matemática. Quizá también la afirmación de Bochenski de que sólo puede compararse con otra obra: con *Los analíticos primeros* de Aristóteles.

Una obra de este valor debería haber comportado una indudable e inmediata repercusión. Sin embargo, a lo largo del breve recorrido, he ido indicando que muchas innovaciones de Frege fueron redescubiertas posteriormente al *Bs*, después de 1879. Lo cual implica la afirmación de que tal influencia directa e inmediata sobre la lógica matemática no existió. Hubo influencia, pero posterior, aunque ciertamente no en el centenario actual; sí a partir de la obra de Russell, fundamentalmente. Ello obliga a unas consideraciones, especialmente referidas al simbolismo creado por Frege.

Es simbolismo que muestra una serie de ventajas indudables para el manejo estrictamente formal. La combinación del condicional con la negación y la concavidad permiten obtener cualquier tipo de expresiones, porque la flexibilidad de este simbolismo es muy superior a cualquier otro. Logra tanto la supresión de paréntesis como mostrar cuál es la estructura de la expresión total; cosa que no consigue la notación polaca, por ejemplo. Sin embargo, también presenta algunas desventajas. Por lo pronto, la ocupación de mucho espacio para las derivaciones, a pesar de todos los proceson abreviadores que se pueden ir creando. Además, no es fácil para la imprenta. Razones, ambas, a las que el mismo Frege replicó en 1896, comparando su Ideografía con la obra de Peano, al indicar que no es ante las dificultades del impresor ante las que el lógico matemático debe rendirse porque "la comodidad del impresor no es ciertamente el *summum bonum*". Igualmente, el hecho de que no siga la escritura tradicional de izquierda a derecha y de arriba abajo, sino que invierta este orden, puede suponer un primer punto de falta de hábito.

No son, sin embargo, objeciones que puedan estimarse suficientemente fuertes, aunque las esgrimiera el propio Russell. No constituyen, desde mi punto de vista, y contra lo que ha venido sosteniéndose, los motivos únicos por los cuales la Ideografía contenida en *Bs*, y continuamente modificada por Frege hasta alcanzar lo que estima su perfección en *Las leyes básicas de la aritmética*, no fuera asumida por el hacer lógico matemático. Hacer que llegó a preferir la notación de Peano y su escuela, aceptada por Russell y Whitehead con algunas modificaciones, o la de Hilbert y su escuela. Veo dos motivos, cuando menos, por los cuales el simbolismo fregeano no fue admitido ni en su época, ni posteriormente, y con el rechazo de su simbolismo se rechazaba el contenido que comportaba.

Por un lado, existe un factor psicológico que enlaza pensamiento y lenguaje y que Frege no tuvo en cuenta. Y si lo tuvo fue, precisamente, para combatirlo. El pensamiento dirigido, conceptual, exige la manipulación mental de palabras como asidero de la construcción y el razonamiento. En el caso matemático, exige de la jerga que permite una mezcla entre signos especiales y palabras del lenguaje ordinario que permiten vehicular el pensamiento. Si éste carece de dicho soporte, tendrá que recurrir al proporcionado por algún tipo de imagen sígnica. Es lo que viene a ofrecer Frege. Pero una imagen despojada de su condición de imagen y, por tanto, inútil para el pensamiento dirigido, conceptual; que dejaría de ser pensamiento dirigido, pensamiento puro en cuanto se dejara arrastrar estrictamente por la imagen sígnica, ya que alcanzaría, en todo caso, no el estado intelectual, sino el de ensoñación.

Por otro lado, fue el propio hecho de que Peano y sobre todo Russell hubieran llegado a las mismas consecuencias de construcción lógica de Frege, y Russell lo hiciera saber así, incluso publicando un elogioso resumen de las ideas del matemático alemán para, a la vez, señalar la aparición de la antinomia que invalidaba el sistema fregeano. Y si esto ocurría en la construcción lógica, se debe tener presente otro punto. Frege había creado la Ideografía con un objetivo: fundamentar la aritmética y no con el puro y exclusivo deseo de crear un sistema lógico en sí. Sistema lógico que no era para Frege más que una pura herramienta. Y en esta labor se había opuesto, por un lado, a la tradición booleana y no sólo en cuanto al simbolismo aritmético (lo cual sería secundario, como mostró Peano con su simbolismo), sino fundamentalmente al punto de partida, a la creencia base de Boole y de prácticamente todos los matemáticos: la axiomatización como elemento constitutivo central del hacer matemático. Por otro lado, Frege, al fundamentar la aritmética, tiene que introducir la noción de clase, aunque pretenda realizarlo con un enfoque intensional y no extensional. Y en este aspecto, Dedekind y Cantor habían logrado no sólo los puntos alcanzados por Frege, sino que habían permitido ir más allá en el hacer estrictamente matemático creando nuevos campos de investigación. Punto en el que fueron ayudados por Peano, quien no se limitó a formular un simbolismo adecuado para realizar una crítica del pensamiento puro, sino que ideó un simbolismo apto y flexible para representar cualquier tipo de expresión matemática, aunque el rigor conceptual fuera menor que el de Frege. Antes de Russell, Dedekind había indicado los paralelos entre su construcción, la cantoriana y la de Frege, con elogio del mismo tipo que el russelliano para este último.

La consecuencia me parece muy clara: no se tenía necesidad de leer a Frege, de hacer el esfuerzo de romper con hábitos muy arraigados de lectura para alcanzar lo ya alcanzado y puesto en notación más clara y legible, más apta para el pensamiento dirigido, conceptual. Incluso hoy, la vuelta a la lectura de Frege se realiza no en función de su Ideografía o de su sistema lógico (que no constituyen en el momento actual más que un mero primer curso inicial de lógica matemática) sino en función de sus ensayos lógico-semánticos, en función de la profundidad de sus análisis respecto a puntos clave de lo que estimar filosofía de la lógica.

Esto último no implica, en modo alguno, que su influencia no haya sido extraordinaria, pero a través, por una parte, de Russell y Russell-Whitehead en *Principia Matemathica*, y por otra, de Carnap y Wittgenstein quienes aceptaron, entre otras cuestiones, la distinción fregeana de la existencia de sólo dos tipos de juicios, analíticos y sintéticos, distinción base para el neopositivismo lógico, así como el criterio de aplicar la lógica matemática para un análisis profundo del lenguaje, tanto del filosófico como del científico. Bien entendido que como aplicación, no como campo de estudio propio. Influencia determinante para lo que vino en estimarse problema filosófico central del siglo XX: la semántica, inconcebible sin la obra de Frege en la forma que adoptó. Influencia que no ha tenido paralelo en el hacer intrínseco matemático, ya que éste siguió la línea cantoriana –línea paralela a la de Frege si se identifica conjunto con extensión de concepto– a la vez que dicho hacer se orientó por la axiomatización formal, orientación opuesta y contraria a las creencias de partida fregeanas. Frege, en este punto, quedó fuera de cauce.

Juegos matemáticos

"Gödel, Escher, Bach", de Douglas Hofstadter

Martin Gardner

Esta frase no verbo.

—DOUGLAS R. HOFSTADTER,
Gödel, Escher, Bach: an Eternal Golden Braid

De cuando en cuando, cada pocos decenios, algún autor desconocido produce una obra de profundidad, alcance, claridad, ingenio y originalidad tales que de inmediato es reconocida la importancia del acontecimiento literario. *Gödel, Escher, Bach: An Eternal Golden Braid,* grueso volumen publicado por Basic Books (777 páginas), es una de estas obras. Su autor (así como ilustrador y maquetista) es Douglas R. Hofstadter, joven especialista en ciencias de computación de la Universidad de Indiana, hijo del renombrado físico Robert Hofstadter.

¿Qué podrían tener en común Kurt Gödel, M.C. Escher y Johann Sebastian Bach? Vemos la respuesta simbolizada en la fotografía de la página opuesta, y en la sobrecubierta del libro. En cada una de ellas se ven dos bloques de madera flotando en el espacio, iluminados de forma que las sombras que proyectan sobre las tres paredes que concurren en un rincón de la sala dibujen las iniciales de los apellidos Gödel, Escher, Bach. Más exactamente, el bloque superior proyecta "*GEB*", encabezamiento de la primera mitad del libro, y el inferior, "*EGB*", título de la segunda mitad. Podríamos imaginar que las letras *G, E, B* son marcas de identificación de tres cabos que se van trenzando al permutar reiteradamente pares de letras. Para completar el ciclo desde *GEB* hasta *GEB* (pasando por *EGB*) se requieren seis pasos.

El Dr. Hofstadter (se doctoró en física por la Universidad de Oregon) llama "tripletas" a estos bloques, condensación de triple letra. Según él mismo explica, la idea se le ocurrió de repente. Su intención inicial era preparar una monografía sobre el teorema de Gödel, pero sus ideas fueron progresivamente cobrando mayores vuelos, incluyendo a Escher y Bach, hasta que finalmente llegó a la conclusión de que las obras de estos hombres no eran sino "sombras proyectadas en distintas direcciones por cierta esencia sólida y central". Se propuso entonces "reconstruir el objeto central", y con ello vio la luz este libro.

Hofstadter talló los bloques de las tripletas en madera de secoya, ayudándose de una sierra de cinta y una fresa de espiga. La idea fundamental que inspira sus formas es evolución de un problema clásico: hallar un cuerpo sólido que proyecte las sombras de un círculo, un cuadrado y un triángulo. ¿Podrán construirse tripletas para todo conjunto de tres letras distintas? Si es lícito distorsionar suficientemente las letras, la respuesta es afirmativa; por tanto, para que el problema ofrezca interés deben imponerse algunas restricciones. Para empezar, las letras (preferiblemente mayúsculas) deben tener todas la forma habitual, y encajar ajustadamente en los tres rectángulos correspondientes a las proyecciones ortogonales de un bloque ortoédrico. Además, el cuerpo buscado debe ser conexo, es decir, no debe descomponerse en varias piezas independientes. Salvo por tanteos, no es fácil determinar si podrá construirse una tripleta así para tres letras dadas. Resulta, en realidad, que algunas de ellas son imposibles. Por otra parte, el problema admite exóticas variantes; por ejemplo, n-tup-letas que den por proyección n letras; tup-letas cuatridimensionales que proyecten tripletas sólidas, que a su vez proyecten sombras planas de letras; cuerpos sólidos tridimensionales que proyecten números, figuras o palabras, etcétera. (Debo esta descripción de las tripletas a Scott Kim, amigo de Hofstadter, y con quien colaboró estrechamente en la preparación de muchas facetas del libro.)

¿Cuál es la realidad que percibe Hofstadter como subyacente a la obra de los tres gigantes? Un aspecto de tal realidad es la estructura formal de la matemática: una estructura que, como pone de manifiesto la famosa demostración de indecidibilidad de Gödel, tiene infinidad de niveles, ninguno de ellos capaz de capturar toda la verdad en un único sistema consistente. Hofstadter lo expresa con suprema tersura: "La noción de demostrabilidad es más débil que la noción de verdad". En todo sistema formal lo bastante rico como para contener la aritmética pueden formularse enunciados verdaderos cuya demostración es imposible dentro del sistema. Para demostrarlos es preciso saltar a un sistema más rico, donde nuevamente pueden formularse enunciados verdaderos imposibles de probar, y así sucesivamente. El proceso continúa indefinidamente.

¿Es gödeliano el universo, en el sentido de que no hay fin para el descubrimiento de sus leyes? Puede que sí. Puede que por hondo que cale la investigación científica haya siempre leyes no captadas por las teorías, una sucesión infinita de ruedas dentro de ruedas. Hofstadter razona elocuentemente en favor de una especie de platonismo en el que la ciencia, en cualquier estadio de su historia, sería a modo de las sombras proyectadas en la pared de la caverna de Platón. La realidad última está siempre más allá de nuestro alcance. Es el Tao, del que nada puede decirse. En el fin del prefacio de su libro, Hofstadter escribe "En cierto modo, este libro es un enunciado de mi religión".

Para legos, no sé de ninguna exposición de los trabajos de Gödel, de su revolucionario descubrimiento y sus consecuencias, que supere a la ofrecida en el libro. El descubrimiento de Gödel se refiere en particular a la recurrencia, la auto-alusión y la regresión infinita; Hofstadter encuentra vívidamente reflejados estos tres temas en el arte de Escher, el más matemático de los artistas gráficos, y en la música de Bach, el más matemático de los grandes compositores. La propia estructura del libro se encuentra tan saturada de complicados contrapuntos como las composiciones de Bach o el *Ulysses* de Joyce. La primera mitad del libro sirve de preludio a la segunda, de la misma forma que un preludio de Bach sirve de introducción a una fuga. Además, cada capítulo está precedido de una especie de preludio, que ya a comienzos del libro toma la forma de "Diálogo" entre Aquiles y la Tortuga. Más tarde hacen aparición otros personajes: el Perezoso, el Oso Hormiguero, el Cangrejo, y finalmente, Alan Turing, Charles Babbage y el propio autor. Cada Diálogo sigue la pauta de una composición de Bach, y en ocasiones la correspondencia es estricta. Otro ejemplo: si la composición tiene n voces, así sucede en el correspondiente Diálogo. Por ejemplo, si alguno de los temas de la composición se repite tocado

de atrás hacia delante, o vuelto del revés, lo mismo sucede en el correspondiente Diálogo. Cada Diálogo enuncia lúdicamente, con juegos verbales increíbles (retruécanos, acrósticos, acrónimos, anagramas y muchos tipos más) los temas que luego serán examinados con mayor seriedad en el correspondiente capítulo.

Hay dos razones principales para que los Diálogos sean iniciados por Aquiles y la Tortuga. Primera, son los primeros actores de la paradoja de Zenón (tema del primer diálogo del libro), donde Aquiles debe alcanzar la Tortuga a pesar de un razonamiento por recurrencia infinita que afirma lo contrario. Segunda, son también los interlocutores de otra paradoja no menos ingeniosa, aunque no tan conocida, ideada por Lewis Carroll. En la paradoja de Carroll, que Hofstadter reproduce en el segundo Diálogo, Aquiles desea demostrar el teorema *Z* de Euclides, partiendo de las premisas *A* y *B*. Sin embargo, la Tortuga no aceptará que el teorema está demostrado mientras Aquiles no postule una regla de inferencia, *C*, que enuncie explícitamente que *Z* es consecuencia de *A* y *B*. Aquiles añade la regla, creyendo de esta forma zanjar la discusión. La Tortuga salta entonces a un nivel superior, exigiendo otra regla de inferencia *D* que enuncie que *Z* se deduce de *A, B* y *C*. La regresión infinita resultante parece invalidar todo razonamiento, de forma muy semejante a como la paradoja de Zenón parece invalidar todo movimiento. "Quedan muchas hojas en blanco, según veo", exclama la Tortuga echando un vistazo al cuaderno de Aquiles. "¡Las necesitaremos TODAS" El guerrero sufre un escalofrío.

Uno de los Diálogos de Hofstadter, titulado "Contracrostipunctus", es un acróstico –en el que no faltan ni los signos de puntuación– que afirma que tomando de atrás hacia delante las palabras que contiene resulta un acróstico de segundo orden que deletrea "J. S. Bach". Otro de los Diálogos, "Crab Canon", ilustrado con una teselación periódica de cangrejos original de M.C. Escher, se basa en el "Crab Canon" de la *Ofrenda Musical* de Bach. Al mismo tiempo que la Tortuga discurre sobre Bach, sus frases se intercalan con las de Aquiles, quien está comentando a Escher. La Tortuga y Aquiles usan las mismas frases, aunque en orden inverso. En el momento del cruce, el Cangrejo hace una breve intervención, anudando los dos cabos del discurso de aquéllos, mitades entretejidas en el tiempo de forma similar a los cangrejos positivos y negativos de la teselación de Escher, que se entretejen en el espacio.

Las iniciales *A, T* y *C* (de Aquiles, Tortuga y Cangrejo) se corresponden con las iniciales de adenina, timina y citosina, tres de los cuatro nucleótidos del ADN, molécula que tiene la extraordinaria capacidad de producir copias de sí misma. Lo mismo que Aquiles forma pareja con la Tortuga, la adenina se empareja con la timina a lo largo de la doble hélice de la molécula de ADN. La citosina se empareja con la guanina. Puesto que la inicial *G* puede tomarse como significando "gen", Hofstadter se vio inducido a efectuar "retoques quirúrgicos en el discurso del Cangrejo", con objeto de que reflejase esta coincidencia. El asombroso paralelismo entre los dogmas de la lógica matemática y el "dogma central" de la biología molecular se muestra vívidamente en una carta que Hofstadter bautiza "Dogmapa central".

La letra *G* puede representar también



"enunciado de Gödel", enunciado que es corazón de la demostración de Godel, y que afirma de sí mismo que es indemostrable. El enunciado proporciona a Hofstadter un modelo de los que denomina "Bucles Raros", y sirve de botón de muestra del concepto de auto-alusión, uno de los temas principales de su obra. Se comentan en ella docenas de bucles raros, desde el canon de Bach, indefinidamente ascendente (que se va modulando en registros más y más agudos, hasta cerrar el bucle, retornando al tono primitivo), hasta curso de agua cerrado consigo mismo de la *Cascada* de Escher, o la escalera que enlaza consigo misma de *Ascenso y Descenso*. Uno de los más divertidos modelos de *G* es un tocadiscos *X* que se destruye a sí mismo cuando se toca con él un disco titulado "No puedo ser reproducido en el tocadiscos *X*".

Ejemplo particularmente llamativo de bucle raro de dos etapas es un dibujo de Escher donde se ven dos manos, cada una de las cuales dibuja a la otra. Nosotros, que vemos la figura, podemos escapar a la paradoja "saliéndonos del sistema", contemplándolo desde un metanivel, de la misma forma que podemos escapar de las paradojas tradicionales de lógica saltando a un metalenguaje. También nosotros poseemos bucles raros, porque la mente humana tiene la capacidad de reflexionar sobre sí misma, es decir, la excitación de neuronas origina pensamientos referentes a neuronas. (El número extraordinario de noviembre de *Investigación y Ciencia* estará dedicado a este tema.) Desde una perspectiva de más altos vuelos, la mente humana se encuentra en un nivel del universo donde la materia ha adquirido la pasmosa capacidad de contemplarse a sí misma.

Hacia el final del libro, Hofstadter ha introducido a sus lectores en la moderna lógica matemática, en las geometrías no euclídeas, en teoría de computabilidad, isomorfismos, enunciados de Henkin (que enuncian su propia demostrabilidad), los postulados de Peano (no despreciando el retruécano "piano"), los diagramas de Feynman para partículas que viajan hacia atrás en el tiempo, el último teorema de Fermat (con su correspondiente broma sobre "fermata"), números transfinitos, la conjetura de Goldberg (hábilmente conectada con las *Variaciones Goldberg* de Bach), máquinas de Turing, ajedrez por ordenador, música computerizada, lenguajes de ordenador (Terry Winograd, especialista en simulación mediante ordenador de lenguajes naturales, aparece en uno de los Diálogos como Dr. Tony Earrwig, anagrama de su nombre), la biología molecular, la "mente" de un hormiguero (en inglés, "anthill") llamado Aunt Hillary, la inteligencia artificial, la consciencia, el libre albedrío, la polémica entre holismo y reduccionismo, y un tipo de enunciados que los filósofos llaman contrafactuales.

Los contrafactuales son asertos fundados en hipótesis contrarias a hechos, como por ejemplo, "Si Lewis Carroll viviese hoy disfrutaría mucho con el libro de Hofstadter". Tales enunciados plantean difíciles problemas de semántica en ciencia, y existe en la actualidad extensa bibliografía relativa a ellos. Para Hofstadter son ejemplos de lo que llama "deslizamiento", es decir, la progresión de sucesos a cosas que son casi reproducción de sucesos. El Diálogo que precede a un capítulo sobre contrafactuales e inteligencia artificial describe un aparato de "Subjun-TV", que permite al observador "repetición instantánea" de

cualquier jugada de un partido de fútbol, y ver cómo se desarrollarían las cosas al modificar ciertos parámetros, por ejemplo, si el balón no fuese esférico, o si estuviese lloviendo, o si el partido se jugase en la Luna o en un espacio de cuatro dimensiones, y así sucesivamente.

También es enormemente interesante la exposición que el libro ofrece sobre el tema de la inteligencia artificial. ¿Obedece el cerebro humano a leyes lógicas? Hofstadter imagina la mente como un sistema jerárquico enmarañado: un sistema de múltiples niveles, dotado de una estructura profundamente auto-alusiva e intrincadamente entretejida. Solamente se ajusta a reglas lógicas en el substrato molecular, "su nivel oculto, formal, de *hardware* ", donde opera con pavoroso silencio y eficacia. En opinión de Hofstadter, ningún ordenador podrá jamás realizar todas las funciones del cerebro en tanto no se consiga reproducir el hardware de éste. En cambio, Hofstadter apenas si conserva la paciencia con el célebre argumento del filósofo anglicano



*Sombras **G, E** y **B** proyectadas por un par de "tripletas"*

J.R. Lucas en el sentido de que el trabajo de Gödel demuestra que un cerebro humano puede pensar por vías imposibles *a priori* para un ordenador.

Tan sólo podemos aquí ofrecer un botón de muestra sobre la faceta recreativa de este libro tan monstruosamente complicado. En el Diálogo que sirve de introducción al análisis de difíciles teoremas de Alonzo Church, Turing, Alfred Tarski y otros, que se titula "The Magnificrab, Indeed" (juegos de palabras sobre el *Magnificat* in *D* de Bach), aparece un extravagante matemático hindú llamado Sr. Najunamar. Najunamar ha demostrado tres teoremas: que puede colorear un mapa de la India con no menos de 1729 colores; que todo número primo par es suma de dos números primos; y ha podido establecer que la ecuación $a^n + b^n = c^n$ no tiene ninguna solución cuando $n = 0$. Los tres son verdaderos, desde luego.

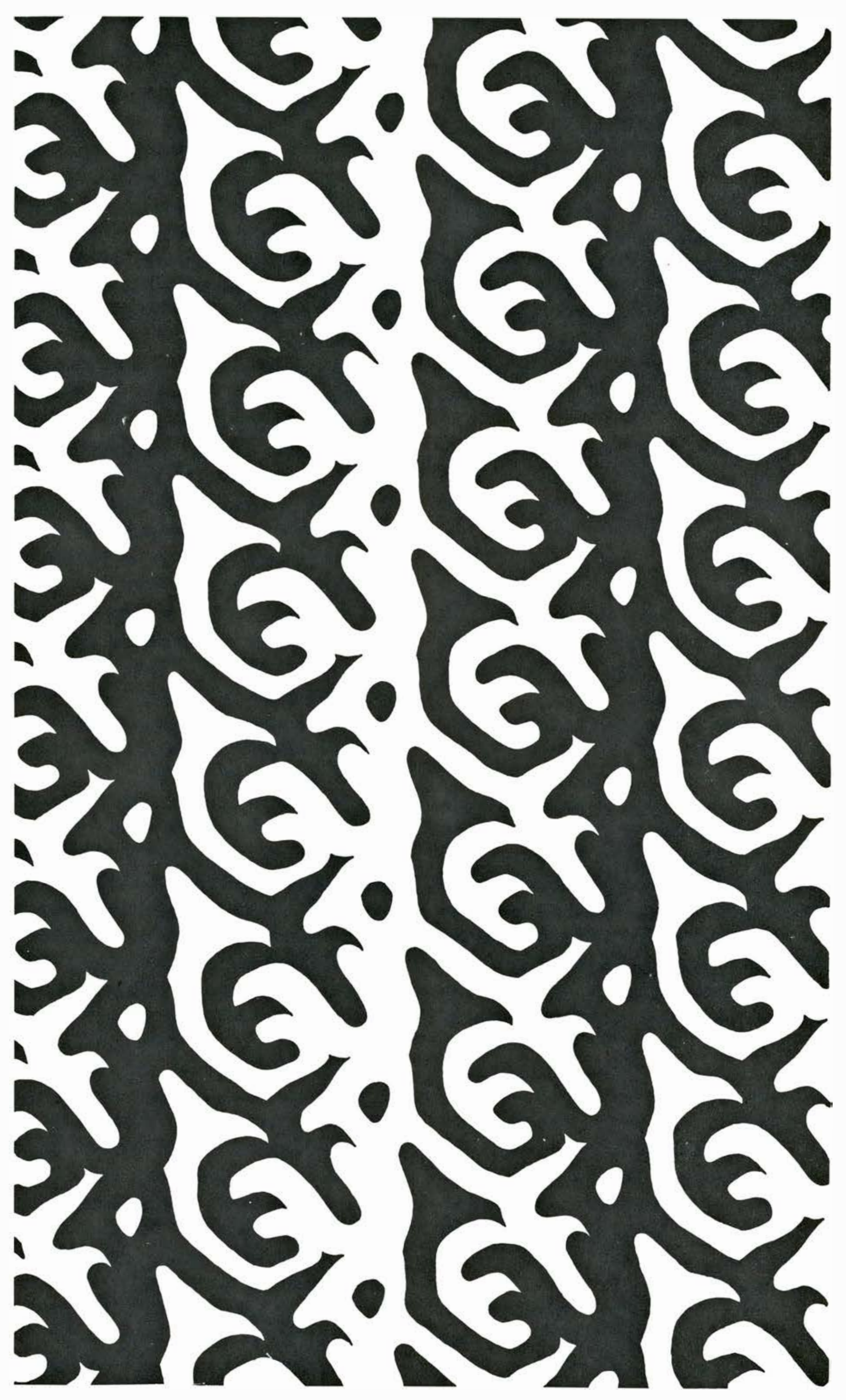
Algunos lectores se habrán dado cuenta de que 1729 es el número del taxi en que viajó G.H. Hardy cuando visitó al matemático hindú Srinivasa Ramanujan ("Najanumar", escrito a la inversa). Hizo Hardy la observación de que el número 1729 era bastante vulgar. Ramanujan replicó instantáneamente que, por el contrario, se distinguía por ser el mínimo entero positivo expresable como suma de dos cubos de un par de formas distintas. Hardy preguntó entonces a su amigo si conocía el número mínimo correspondiente a cuartas potencias. Ramanujan no conocía este número, aunque conjeturaba que sería bastante grande. Hofstadter menciona la solución: 635.318.657, que es igual a $134^4 + 133^4$ y también a $158^4 + 59^4$. Se pregunta entonces si los lectores podrán encontrar el mínimo número expresable como suma de dos cuadrados de dos formas distintas, aunque él no da la respuesta. ¿Podrá el lector determinarla antes de que la demos aquí el mes que viene?

Para explicar el significado del concepto "sistema formal" Hofstadter abre su libro con un sencillo ejemplo que solamente utiliza los símbolos *M, I* y *U*. Estos símbolos pueden disponerse en hileras llamadas teoremas, de acuerdo con las reglas siguientes:

1. Si la última letra de un teorema es *I*, puede añadirse *U* al teorema.
2. A todo teorema *Mx* puede añadirse *x*. (Por ejemplo, *MUMUM* resulta de *MUM; MU* puede transformarse en *MUU*).
3. Si *III* está formando parte de un teorema puede reemplazarse por *U*, aunque la operación recíproca no es aceptable. (Por ejemplo, *MIII* puede transformarse en *MU*, y *UMIIIMU* puede transformarse en *UMUMU*.)
4. Si *UU* está formando parte de un teorema, puede suprimirse. (Por ejemplo, *UUU* puede transformarse en *U*, y *MUUUIII* puede ser transformado en *MUIII*.)

El sistema tiene solamente un "axioma": en la formación de teoremas es obligatorio empezar por *MI*. Toda hilera que pueda formarse por aplicación de las reglas, en cualquier orden, es un teorema del sistema. Así pues, *MUIIU* es un teorema, porque puede generarse a partir de *MI* en seis pasos. Jugando un poco con el sistema *M, I* y *U*, construyendo teoremas al azar, pronto se descubrirá que todos los teoremas empiezan por *M*, y que la *M* no puede presentarse en ningún otro lugar.

He aquí un problema: ¿Es *MU* un teorema? Daré la respuesta de Hofstadter el mes que viene. No diré aquí nada más acerca de *MU*, salvo que tiene en el libro



La **figura** ***FIGURE-FIGURE*** *de Scott Kim*

muchos otros papeles; en particular son las dos primeras letras de "Mumon", un monje Zen que aparece en un delicioso capítulo dedicado a los koans del Zen.

Incluso un sistema tan sencillo como el formado por *M, I* y *U* le permite a Hofstadter suscitar una profunda cuestión. Si de todas las posibles hileras del sistema separamos las que formen teoremas, nos restarán las hileras que no lo sean. Por tanto, la "figura" (el conjunto de teoremas) y el "fondo" comprendido entre los teoremas (o sea, el conjunto de no-teoremas) parecen contener información equivalente. ¿Sucede efectivamente así? ¿Es análogo este sistema a una teselación de Escher, donde los huecos entre animales de una especie son animales de otra, por lo que al reproducir las siluetas de cualquiera de los dos conjuntos quedan automáticamente definidas las del otro? (O bien, ¿es lo mismo una cebra blanca con listas negras que una cebra negra con listas blancas?). A este respecto, Hofstadter reproduce una notable teselación de Kim, donde la palabra "FIGURE" se repite periódicamente en negro, de manera tal que el fondo blanco entre las letras negras forma la misma palabra.

Esta misma idea está festivamente ilustrada en el Diálogo "Sonata para Aquiles solista" (cuyos modelos son las sonatas para violín solo, de Bach). En ella podemos solamente oír el lado Aquiles de una conversación telefónica con la Tortuga, que trata del tema "figura" y "fondo". A partir de la mitad de la conversación podemos reconstruir las frases de la Tortuga.

Los números naturales proporcionan otros ejemplos del esquema "figura-fondo". Podemos determinar todos los números compuestos sin más que suprimir del conjunto de enteros positivos todos los números primos. ¿Sucederá lo mismo en todo sistema formal? ¿Podremos siempre extraer todos los teoremas del conjunto de todos los enunciados posibles en el sistema, y hallar que el residuo —la colección de no-teoremas— es otro sistema formal complementario del primero? Uno de los resultados menos esperados de la moderna teoría de conjuntos es que no siempre ocurre así. Con terminología más técnica: existen conjuntos enumerables por recurrencia que no son recursivos. De esta forma, Hofstadter lleva a sus lectores desde consideraciones casi triviales hasta algunas de las más profundas regiones de la moderna matemática.

El libro concluye con el frenético Diálogo "Ricercar en seis partes", que sigue a un tiempo la pauta del ricercar en seis partes de Bach, y la historia de cómo llegó éste a componer su *Ofrenda Musical.* (Un ricercar es una fuga de tipo muy complicado.) En este Diálogo, Turing y Babbage, pioneros de los ordenadores, van improvisando, sentados al teclado de un flexible ordenador llamado "idiota inteligente", que puede ser tan torpe o tan sagaz como su programador desee. (El nombre del ordenador se inspira en "pianoforte", que significa "tocar suave-tocar fuerte".) Turing produce en la pantalla de su ordenador una simulación de Babbage. Sin embargo, Babbage aparece en ella mirando la pantalla de su propio "idiota inteligente", sobre la cual ha conjurado una simulación de Turing. Cada uno de estos personajes insiste en decir que él es real, y que el otro no es más que un programa del ordenador. Se intenta entonces resolver el debate mediante el Juego de Turing, que fue ideado por Turing como posible método para distinguir un ser humano de un ordenador programado, y que consiste en formular hábiles preguntas. La conversación de esta escena es parodia de la dada por Turing en su clásico artículo sobre el tema.

Llegado este punto, el propio Hofstadter entra en escena y convence a Turing, Babbage y los demás de que no son más que criaturas de su imaginación. El mismo, sin embargo, es tan irreal como los demás personajes del Diálogo, pues también él ha sido imaginado por el autor. La situación recuerda un cuadro de René Magritte titulado *Los dos misterios,* donde la figura de una pequeña pipa aparece con una nota que dice (traduciendo del francés) "Esto no es una pipa". Flotando por encima de la falsa pipa se encuentra otra mayor, presumiblemente auténtica, aunque desde luego, también pintada en el lienzo.

¿Y hasta qué punto era real Magritte? ¿Hasta qué punto somos reales Hofstadter, usted y yo? ¿No seremos acaso sino las sombras de las páginas del libro que llamamos Universo, que tan rápidamente van pasando? Hemos retornado a un platonismo gödeliano, donde la realidad está organizada en infinidad de niveles. ¿Quién podrá decir lo que en verdad es la realidad? La última palabra del libro, "RICERCAR", es un juego de palabras de múltiples niveles, del que son precursores una serie de acrónimos, como la observación de Turing, "Rigid Internal Codes Exclusively Rule Computers and Robots" (Los códigos internos rígidos gobiernan exclusivamente robots y ordenadores), lo mismo que en la inscripción de Bach "Regis Iustu Cantio et Reliqua Canonica Arte Resoluta", escrita sobre una página musical que envió a Federico el Grande. El ricercar en seis partes de Bach está tomado de su *Ofrenda Musical,* con cuya historia comienza el libro. De esta forma, RICERCAR sirve, de una manera que resulta muy parecida a "riverrun", la primera palabra de *Finnegans Wake,* para retorcer la obra en un gigantesco bucle autoalusivo.

Uno de los más divertidos ejemplos de Bucle Raro se encuentra en el Diálogo que prologa el capítulo 16. Allí, el Cangrejo menciona su intención de hojear un libro "chiflado" sobre "lógica metálica" titulado *Copper, Silver, Gold: An Indestructible Metallic Alloy*. En la bibliografía comentada que da Hofstadter puede verse que este libro ha sido escrito por un tal Egbert B. Gebstadter (obsérvense el grupo *EGB* de Egbert, el *GEB* de Gebstadter y las iniciales *EBG*) y fue publicado por Acidic Books (el editor de Hofstadter es Basic Books). He aquí el comentario de Hofstadter: "Formidable batiburrillo, atiborrado y confuso; no obstante, en muchos aspectos notablemente similar a la obra presente. Contiene algunos excelentes ejemplos de auto-alusión indirecta. Tiene especial interés una referencia de su bien comentada bibliografía relativa a una obra isomorfa, aunque imaginaria".

He aquí las soluciones de los problemas de ajedrez del mes pasado.

El problema de Sam Loyd, de jugadas simétricas, en que las Blancas dan mate en el cuarto movimiento tiene dos soluciones:

	Blancas	Negras
(1)	1.P4AD	P4AD
	2.D4T	D4T
	3.D6A	D6A
	4.DxA (mate)	
(2)	1.P4D	P4D
	2.D3D	D3D
	3.D3TR (o 5A)	D3TR (o 5A)
	4.DxA (mate)	

La partida de jugadas simétricas en que las Blancas se dan mate a sí mismas en la octava jugada se desarrolla así:

Blancas	Negras
1.P4R	P4R
2.R2R	R2R
3.R3R	R3R
4.D3A	D3A
5.C2R	C2R
6.P3CD	P3CD
7.A3T	A3T
8.C4D (jaque)	

La única jugada posible para las Negras es PxC, que da jaque mate al rey blanco.

Para resolver el chascarrillo sobre el mate del último problema propuesto el mes pasado, basta con que el Blanco levante su caballo un dedo o dos sobre su casilla, y exclame "¡Mate descubierto".

Taller y laboratorio

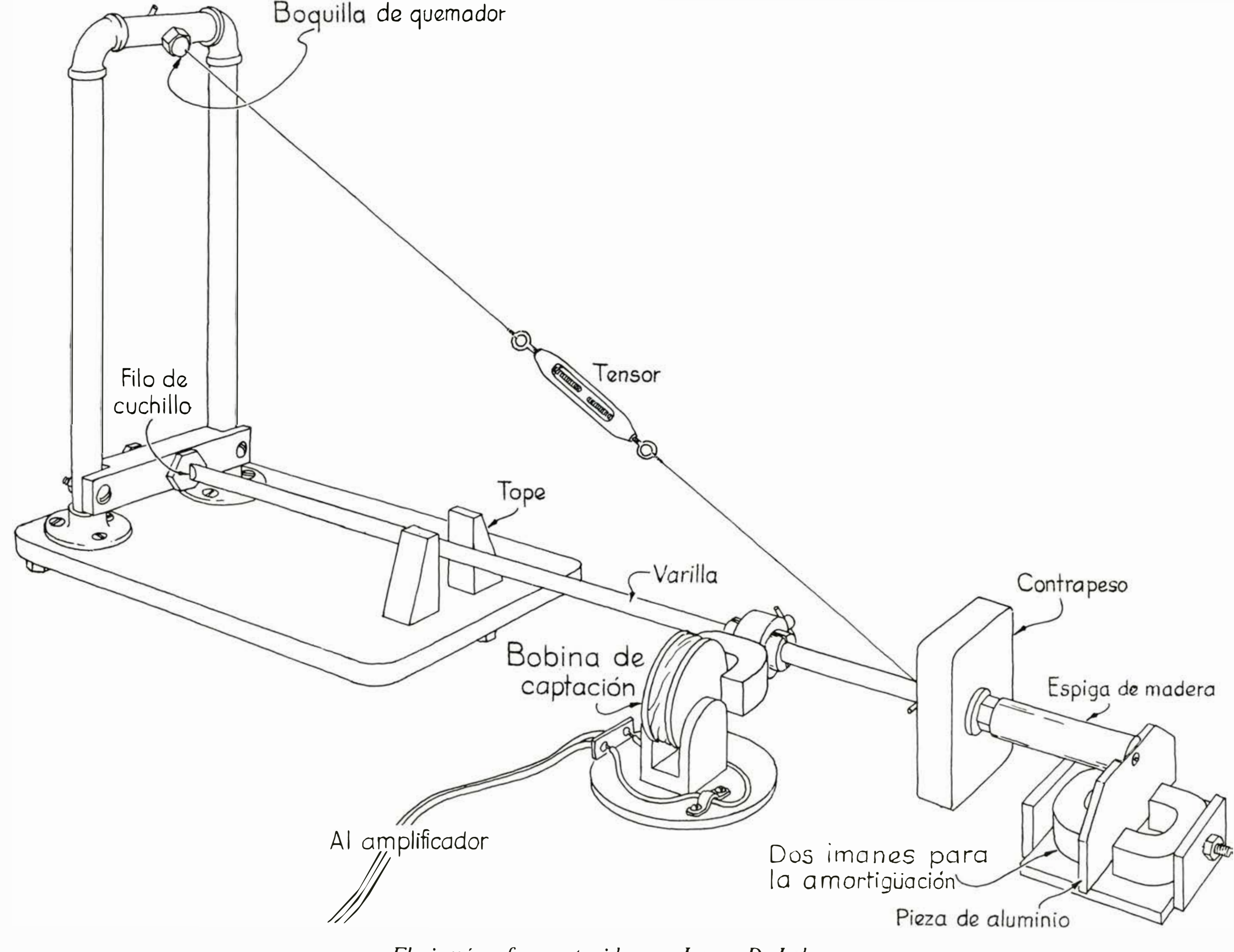
Cómo construir un sencillo sismógrafo para registrar en casa los temblores de tierra

Jearl Walker

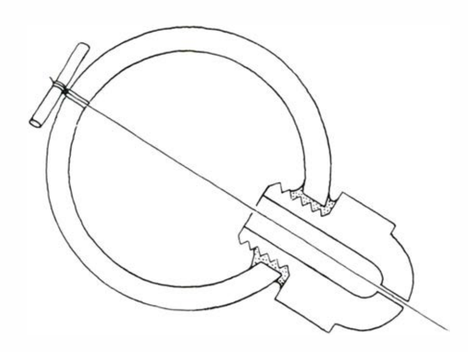
Este mes describiremos el sismógrafo construido por James D. Lehman, de la Universidad James Madison en Harrisonburg, Virginia. El propio Lehman me ha enviado las instrucciones para construir el aparato. Además de poderse fabricar con bastante facilidad, es lo suficientemente sensible como para registrar los terremotos que se produzcan a unos 3000 kilómetros de distancia y tengan una intensidad igual o mayor que 4,8 en la escala de Richter y también todos los terremotos de intensidad 6, como mínimo, originados en cualquier lugar de la tierra. Lehman registró una prueba subterránea de un artefacto nuclear llevada a cabo en Nevada así como un intenso terremoto ocurrido en Turquía.

El principal objetivo de cualquier sismógrafo es transformar un ligero movimiento de la tierra en otro movimiento mucho mayor que pueda ser grabado sobre un soporte para su examen posterior. Las ondas procedentes de un terremoto ocurrido a una gran distancia no llegan todas al mismo tiempo; unas toman trayectorias más largas a través de la tierra y otras tienen una velocidad menor. Por tanto, el registro debe dibujarse en función del tiempo, de manera que la llegada de los varios componentes de las ondas pueda ser anotada y cronometrada. Los intervalos transcurridos entre los diferentes tiempos de llegada de los distintos tipos de ondas pueden ser utilizados para calcular la distancia a la que se ha producido el terremoto.

De los diversos diseños posibles de sismógrafos, Lehman ha elegido uno de tipo voladizo unido a una tira de papel que servirá de registro. Su aparato consiste básicamente en una varilla con un contrapeso, un imán en forma de herradura incorporado a la varilla y una bobina de alambre situada entre los dos polos del imán. Cuando las ondas sísmicas llegan al sismógrafo, el imán y la bobina se mueven uno con respecto al otro, con lo cual se varía el flujo magnético a través de la bobina. El voltaje generado en la bobina por la variación del flujo se amplía y registra sobre la cinta de papel. Un ligero movimiento del suelo termina, por tanto, transformándose en una gran oscilación sobre el registro. Como el papel se mueve delante del registrador a una velocidad concreta, puede determinarse el inicio y la dura-



El sismógrafo construido por James D. Lehman



Boquilla para el cable de soporte

ción del movimiento a partir de la gráfica obtenida.

El extremo de la varilla más alejado del contrapeso termina de forma afilada y se mantiene apoyado en un bastidor fijo. El otro extremo está unido a la parte superior del bastidor por un alambre. Los dos puntos de apoyo de la varilla (el extremo afilado y el punto donde el alambre está unido al bastidor) no se hallan en línea vertical. Por tanto, cuando la varilla comience a oscilar inducida por una onda sísmica, lo hará horizontalmente con un movimiento de tipo pendular alrededor de su posición de equilibrio. Caso de tener una oscilación libre de este tipo, lo único que nos mostraría el registro serían las naturales oscilaciones de una varilla movida por las ondas sísmicas, pero se perdería la información sobre los períodos y horas de llegada de cada una de las ondas.

Para evitar dicha pérdida, Lehman ha diseñado su sismógrafo de forma que amortigüe o elimine las oscilaciones rápidas de la varilla. Este amortiguamiento se lleva a cabo gracias a una pieza de aluminio montada en el extremo de la varilla de manera que oscila dentro de un campo magnético. Este movimiento genera una corriente eléctrica en el aluminio, la cual genera, a su vez, su propio campo magnético. Estos

Tuerca pegada al travesaño

Filo de cuchillo

Varilla

Cabeza pulida y parte del eje roscado de un perno de máquina

Colocación del extremo afilado

campos magnéticos inducidos se oponen a nuevos movimientos de la varilla, que pronto termina sus oscilaciones.

Si se detiene la oscilación natural, el sismógrafo registra la velocidad relativa entre la varilla y el suelo, toda vez que la varilla y la bobina registradora se mueven una respecto a la otra en razón de sus distintas masas de inercia. La corriente creada en la bobina es proporcional a la velocidad con que dicho elemento se mueve.

Si no hubiera amortiguamiento, la varilla sería más sensible a las ondas sísmicas que tuvieran el mismo período que el del movimiento pendular natural de la propia varilla. Con el amortiguamiento la varilla será más sensible a las ondas sísmicas cuyos períodos sean aproximadamente la mitad de la longitud del período pendular natural de la varilla. Un instrumento construido para responder ante ondas sísmicas con períodos de aproximadamente un segundo se llama sismógrafo de período corto; el que responde a períodos de 10 a 20 segundos es un sismógrafo de período largo. Lehman ajustó su varilla para que su período natural fuera de 12 a 18 segundos; construyó, pues, un instrumento de período largo.

El bastidor principal del sismógrafo consta de una base y dos secciones verticales de tubería. Se aconseja emplear una base metálica rígida, pero también puede utilizarse madera curvada o bien un tablero de conglomerado con suficiente espesor; en este último caso hay que contar con las deformaciones que puede provocar la humedad del ambiente. Lehman construyó su base con una chapa metálica de 40 por 25 centímetros, afirmada con tres tornillos de un centímetro, de los que soldó dos en la parte de atrás y uno en el centro de la parte delantera. Estos tres puntos de apoyo hacían que la base estuviera libre de cualquier sacudida. Pueden colocarse plaquetas metálicas plegadas debajo de los soportes de la base para darle la inclinación deseada antes de poner en funcionamiento el sismógrafo.

La varilla, de acero, mide unos 8 milímetros de diámetro por 75 centímetros de longitud. Uno de los extremos está afilado. En el otro se ha practicado una rosca de unos ocho centímetros. Por dicha rosca se introduce una tuerca hasta el final y, a continuación, un trozo de plomo de 2 a 3 kilogramos de peso, fijándolo con una segunda tuerca colocada acto seguido.

La parte vertical del sismógrafo se halla unida a la base por medio de dos bridas de tubería de 12 milímetros. Ambas están atornilladas a la base de tal manera que sus bordes queden aproximadamente a centímetro y medio de la arista de dicha base. Para unir las bridas a ésta, pueden usarse pernos grandes de madera o bien tornillos metálicos de 6 milímetros. Hay que procurar que dichos tornillos no obstruyan los que sirven de apoyo. Y hasta puede que hayamos de pulir las cabezas de los tornillos de sujeción.

Se introducen dos trozos de tubería de 12 milímetros, con una longitud igual de 45 centímetros y con los extremos roscados, enroscándolos en las bridas. Se situarán dos travesaños entre las piezas verticales, en dos sitios distintos: uno donde el extremo afilado de la varilla choca contra la estructura y otro donde se sujete el alambre que soporta el otro extremo de la varilla. El travesaño inferior es una plaquita de metal atornillada a los tubos verticales. El travesaño superior es un trozo de tubería de 10 centímetros de longitud que se une a las piezas verticales por medio de dos codos con rosca de 12 milímetros.

El alambre que soporta a la varilla se sujeta al centro del travesaño superior. Lehman usa, como punto de sujeción del alambre, la boquilla de un quemador. Las boquillas típicas suelen tener un diámetro de 14 milímetros. Hay que quitar el filtro interior ya que es necesario que el alambre entre por el diminuto agujero. Montar la base, las bridas, las tuberías verticales y el travesaño superior para determinar dónde debe sujetarse el alambre de la varilla. Hay que procurar que el punto de sujeción quede más o menos en el centro del travesaño superior, de suerte que, cuando el aparato quede montado, el alambre forme un ángulo de 30 a 40 grados con la horizontal. Luego se hará una marca en el sitio adecuado sobre la tubería horizontal y se pasará por allí la tubería con un agujero de 14 milímetros. Introducir en dicho agujero la boquilla y meter por ella el alambre que habremos de sacar por el otro lado de la tubería, donde se anclará.

Lehman sugiere que se emplee bien hilo 26 Nichrome o bien una cuerda de piano. El hilo se dividirá en dos partes, atándose un extremo de cada una de ellas a la boquilla y a un agujero que se habrá practicado en el extremo más alejado de la varilla, respectivamente, en tanto que los otros dos extremos quedan enfrentados entre sí a través de un tensor, con el que haremos que el hilo o alambre quede tirante.

El travesaño inferior es un trozo plano de metal atornillado a los trozos de tubería verticales, a unos cinco centímetros de la base. A él debe engarzarse un trozo de metal duro, como la cabeza de tornillo pulida. El borde afilado de la

varilla se apoyará sobre dicha plaquita de metal duro con el borde afilado en posición vertical. Cuando el sismógrafo esté por fin montado, la boquilla debe sobresalir como un centímetro de la vertical de la placa donde apoya el borde afilado de la misma. A tal fin resultaría provechoso colocar una superficie metálica ajustable.

Incorpórese luego a la varilla un imán en forma de herradura. Lehman tiene uno con 25 kilogramos de fuerza tractora. El imán deberá tener una abertura de 2,2 centímetros (7,8 pulgadas). En dicha abertura deberá colocarse la bobina que enviará al equipo eléctrico las señales referentes al movimiento relativo entre el imán y la bobina. Lehman usa una abrazadera y unos trozos de madera para unir el imán a la varilla cerca del fragmento de plomo.

Tendremos la bobina registradora si damos 10.000 o más vueltas de hilo magnético número 34 en un carrete. Si se posee un imán de Edmund Scientific la bobina debe tener 1,9 centímetros de anchura y 8 de diámetro para que pueda colocarse con holgura de 5 a 10 milímetros en el interior de la abertura del imán, y disponga así suficiente espacio libre por ambos lados. Lehman se construyó su bobina o carrete pegando dos piezas circulares de plástico de 2 milímetros de espesor a un cilindro de madera de 2,5 centímetros de diámetro. Colocó el carrete en el eje de una máquina taladradora y fue enrollando cuidadosamente el hilo mientras hacía girar la máquina a baja velocidad (hay que prestar mucha atención al manejo de la máquina, ya que no es difícil cortarse en un dedo con el hilo). A través de un cable eléctrico normal llevó los extremos del hilo al amplificador. El cable no debe medir más de tres metros de longitud, pues de lo contrario nos toparíamos con una resistencia eléctrica demasiado grande.

Pondremos la bobina sobre una base estable, con tres puntos de apoyo como la base principal. Nos aseguraremos de no emplear ningún material magnetizable. Así, para la base se recurre a una pieza de aluminio con soporte de latón para que la bobina opere a mayor celeridad.

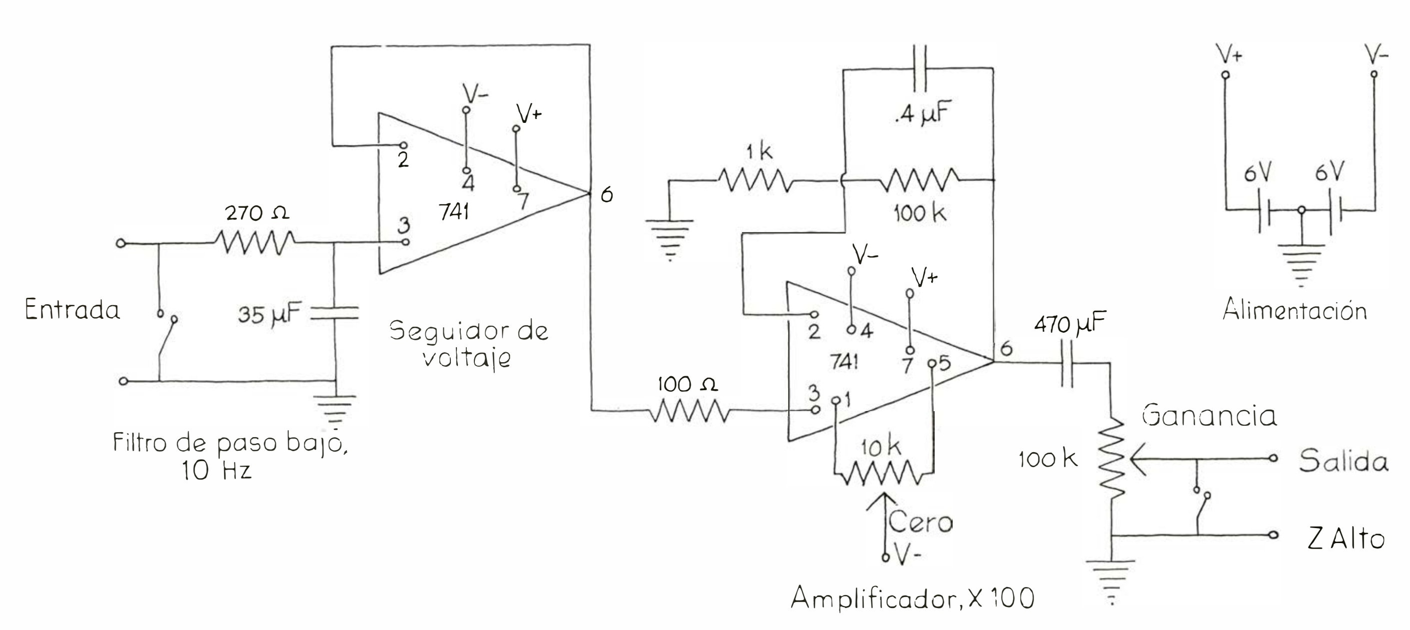
En el otro extremo de la varilla, el no afilado, se necesita otro campo magnético que amortigüe las oscilaciones de la varilla. Lehman montó dos imanes Edmund sobre una base de madera, orientándolos de forma que se atrajeran mutuamente. Después enroscó una clavija de madera al extremo de la varilla, y al extremo de la clavija fijó una pieza plana de aluminio (8 centímetros de ancho) con la idea de que dicha pieza quedara situada entre los dos imanes, de manera que, cuando la varilla oscile, la pieza de aluminio también lo haga entre ambos imanes. La corriente que entonces se creará en el aluminio generará un campo magnético que se opondrá al movimiento, frenando por tanto las oscilaciones de la varilla.

El sismógrafo debe estar totalmente cubierto al objeto de protegerlo de las corrientes de aire que se producen en la habitación, debidas tanto a la convección del lugar como al aire acondicionado. La montura de protección será de cristal o de plástico transparente para que pueda observarse.

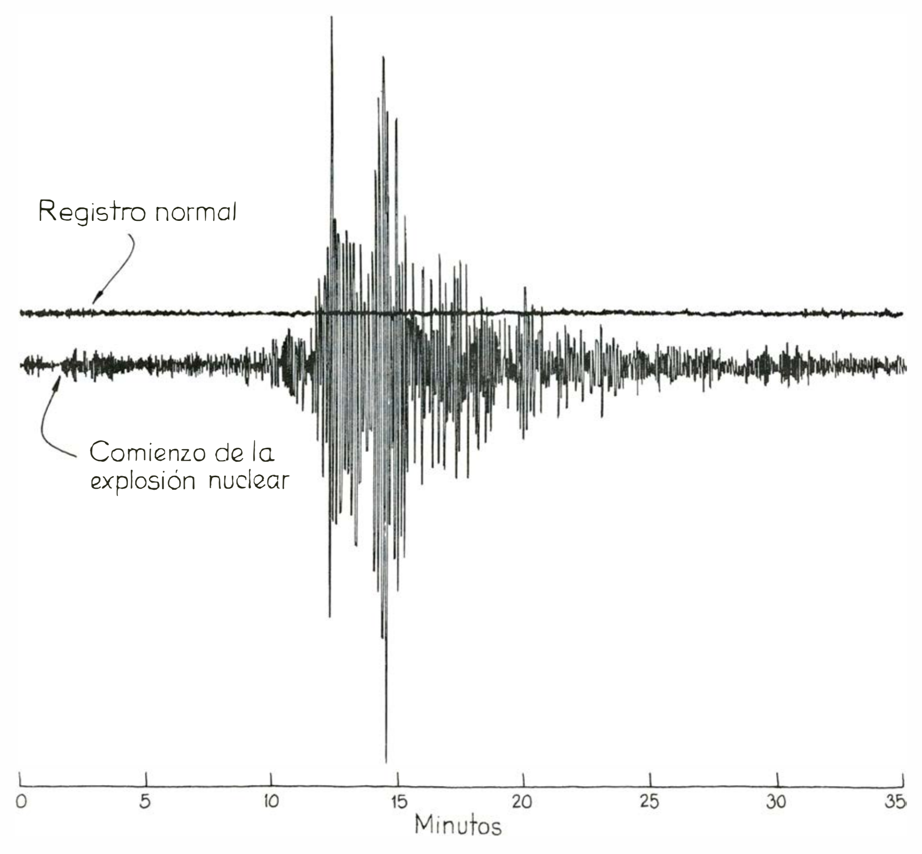
El voltaje de la bobina se leerá a través del amplificador. Podemos comprar un amplificador de corriente continua o bien nos lo podemos fabricar siguiendo las indicaciones de la ilustración inferior. Lehman dice que dicho circuito opera normalmente multiplicando la señal por un factor de 10. El circuito se puede montar sobre un tablero de fibra o bien dentro de una pequeña caja. Las interferencias a 60 hertz procedentes de la instalación eléctrica de la habitación se filtran en la etapa de entrada al amplificador así como cualquier ruido mecánico en un rango de frecuencia de 10 hertz o más. Las señales sísmicas caen dentro del rango de 1 a 1/15 hertz y, por tanto, pasan a través del filtro. La primera entrada 741 actúa como un seguidor de voltaje con respecto al voltaje de la bobina registradora. La segunda entrada 741 actúa como un centumultiplicador. Lehman mantiene que con dos pilas, tipo linterna, de seis volt, el amplificador puede trabajar durante varios meses.

La salida del amplificador ha sido diseñada para ajustarse al nivel de los 10 milivolt del registrador gráfico de Lehman. Comenta que la velocidad de paso del papel de ese modelo es tan precisa que no necesita un sistema adicional de control de tiempo. Para la impresión se pueden utilizar de forma satisfactoria rotuladores. Los cartuchos de tinta pueden volverse a utilizar rellenándolos con un cuentagotas de los que se expenden en farmacias. Durante un día de marcha se gastan aproximadamente dos tercios de un cartucho.

Si se tiene que comprar otro tipo de registrador gráfico, Lehman pone mucho énfasis en la importancia de encontrar uno que controle de forma independiente la velocidad del papel y el movimiento de la pluma. El sujetador de la



Circuito del amplificador



Registro de Lehman de un experimento nuclear

pluma debe ser articulado de forma universal y el papel fácilmente extraíble rebobinando o por cualquier otro sistema.

Lehman consiguió papel de 25 centímetros standard para su aparato y lo hacía pasar a una velocidad de 30 centímetros por hora, de manera que los intervalos en el papel eran de medio centímetro por minuto. Al aprovechar el papel por ambos lados Lehman reduce el coste del mismo a la mitad. Calculó que con impresión diaria durante todo el año el coste de papel sería de sólo 10 dólares.

Para poner en funcionamiento el sismógrafo hay que quitarle la varilla y el cable de soporte, después se nivela la base colocando tantas arandelas de metal o cuñas debajo de las patas cuantas sean necesarias. Para nivelarlo bien se aconseja utilizar un nivel de burbuja. La boquilla del soporte superior debe estar aproximadamente un centímetro por fuera de la vertical donde se apoyará el extremo afilado de la varilla en el soporte inferior. Para conseguir esto es posible que sea necesario cambiar la pieza de soporte inferior o poner arandelas bajo las dos patas traseras de la base.

Se coloca luego la varilla. Se ata el cable al tensor de modo que la varilla esté horizontal y el borde afilado esté vertical en su placa. Para comprobar el período de oscilación hay que empujar ligeramente la varilla y el período debe durar entre cuatro y ocho segundos. Hay que comprobar también su posición en reposo, siendo la normal horizontal y perpendicular a la placa de metal sobre la cual descansa el extremo afilado. En el otro extremo de la varilla pueden añadirse topes a ambos lados, a unos 6 milímetros de distancia de la posición de reposo del péndulo, para evitar que éste oscile mucho mientras se está ajustando el sismógrafo. Con los topes bien colocados la posición de reposo debería estar obviamente a mitad de camino entre ambos.

Se debe dejar el sismógrafo inactivo durante varios días, al objeto de que el marco se ajuste a la tensión y el cable se estire. Se ajustará luego la varilla de forma que su período natural de oscilación sea de 12 a 18 segundos. Para elevar el período de oscilación a esta frecuencia hay que poner una placa de metal debajo de la ya existente donde se apoya el extremo afilado o bien el soporte delantero de la base. De esta forma se reduce la distancia en que la boquilla se aparta de la vertical del punto de apoyo del borde afilado. Si se quisiera inclinar el conjunto tanto que la boquilla estuviera justo en la vertical del filo de cuchillo, la varilla ya no oscilaría como un péndulo normal, porque no habría fuerza recuperadora que lo impulsara a retornar a su posición de reposo. Cuando el conjunto está debidamente alineado, la boquilla se halla fuera de la vertical del extremo afilado y la varilla tiende constantemente a su posición de reposo por su propio peso. Cuanto más lejos esté la boquilla de la vertical referida, mayor será la fuerza recuperadora y menor el período de oscilación natural de la varilla. Hay que ajustar la alineación de la boquilla y el extremo afilado, de modo que el período natural sea entre 12 y 18 segundos. La varilla se mostrará así sensible a las ondas sísmicas de períodos ligeramente más cortos.

El siguiente paso va a ser ajustar la amortiguación. En primer lugar se comprueba ésta colocando los imanes de amortiguación en su sitio alrededor de la placa vertical de aluminio y, luego, haciendo oscilar el extremo de la varilla sobre una distancia de un centímetro aproximadamente. Cuando se suelta el extremo, oscilará hasta sobrepasar la posición de descanso. Con la amortiguación adecuada, el exceso deberá ser de unos dos milímetros. Se puede ajustar la amortiguación por medio de la ranura existente entre los imanes; si se cierra la ranura se incrementa el campo magnético y la amortiguación. También se puede ajustar haciendo variar el período de oscilación natural de la varilla. Así, un período más largo significaría menor fuerza restauradora, de modo que con el mismo campo magnético la amortiguación se incrementaría.

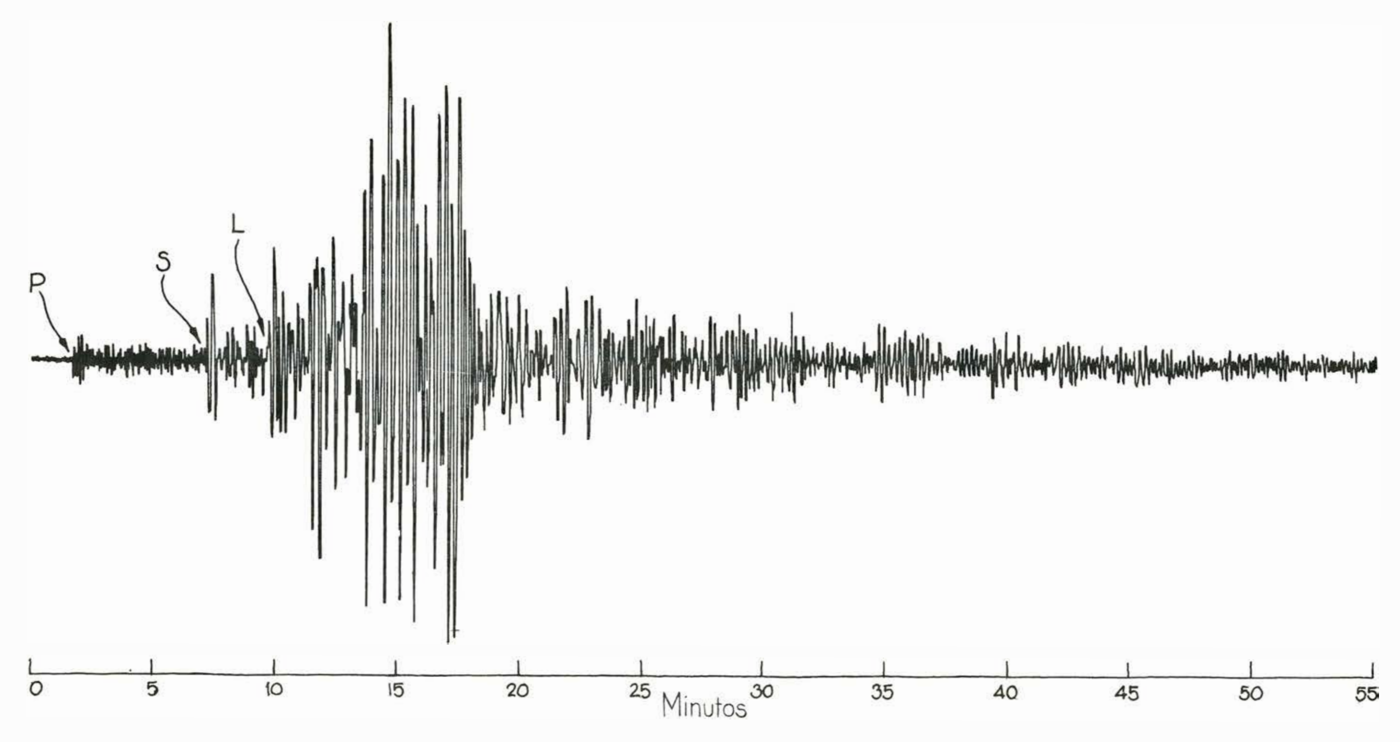
Con este proceso se consigue un ajuste grosero de la amortiguación. Sitúese luego la bobina en relación al imán que está unido a la varilla, de suerte que ambos estén a la altura adecuada del suelo y que la bobina se halle a unos 6 milímetros dentro de la ranura del imán cuando el fiel señala su punto de reposo. Se cubre todo el conjunto. Ajustamos la ganancia del amplificador hasta que pueda verse el ruido de fondo registrado en el papel. Lehman dice que si entonces se acerca uno al sismógrafo la vibración del suelo producida por el peso debería registrarse en el papel del mismo con una amplitud de unos 7 centímetros.

Para comprobar la amortiguación hay que acercarse andando hasta el sismógrafo, esperar unos segundos y después alejarse. El papel del registro deberá mostrar dos picos por el avance y dos por el retroceso. El primer pico de cada par corresponde a la vibración del suelo al avanzar; el segundo, más pequeño, representa el exceso de la varilla al sobrepasar su posición de reposo. Si la varilla está debidamente amortiguada, la razón de los dos picos debería oscilar entre 6:1 y 10:1; en caso de que no ocurra tal, habrá que ajustar la amortiguación. Lehman señala que es necesario tener

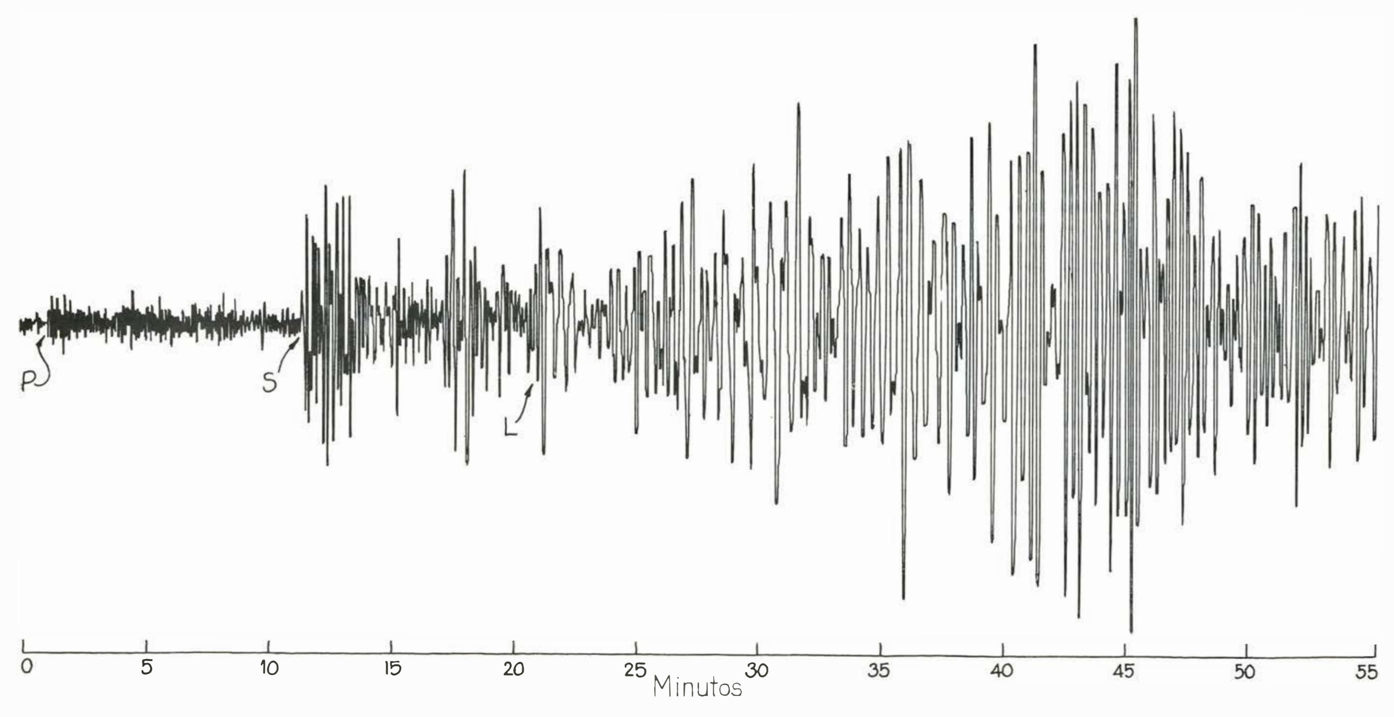
paciencia hasta alcanzar la amortiguación correcta, pero una vez que se consigue, el instrumento no necesitará mayor atención durante semanas o incluso meses.

El mejor sitio para colocar el sismógrafo es una habitación donde no entre nadie y donde no se produzcan cambios térmicos. No hay que colocarlo donde le dé el sol ni en el suelo, porque las variaciones resultantes en la tensión del suelo se registrarán en las lecturas. Se debe tener en cuenta que, aun cuando el sol no entre directamente en la habitación, ésta puede sufrir cambios térmicos causados por los rayos de sol que dan en la fachada. Lehman tuvo éxito colocando uno de sus aparatos en un suelo de arena en una parte inacabada de su patio. Dado que el sismógrafo es más sensible a las ondas que pasan perpendicularmente a la varilla, y puesto que la mayoría de los terremotos se producen en un azimut este-oeste, convendrá instalar el sismógrafo con la varilla en dirección norte-sur.

Cuando se empiezan a hacer registros se aprecia un ruido de fondo constante en las lecturas; gran parte de este ruido se debe a microseísmos, que son pequeñas oscilaciones de la tierra que aún no han sido debidamente explicadas. Los investigadores han intentado desde hace mucho tiempo relacionar los microseísmos con fenómenos naturales tales como los golpes del oleaje en las costas acantiladas o la presencia de un ciclón en el océano. Aunque algunas veces la conexión entre los microseísmos y tales



Registro de un terremoto en Las Galápagos el 29 de marzo de 1976



Registro resultante de un terremoto en Turquía el 24 de noviembre de 1976

causas ha parecido muy clara, en otras se decía que no tienen relación alguna.

En algunos de los datos que Lehman me ha enviado se pueden identificar nuevas causas de ruidos en los registros. Por ejemplo, un día un pequeño grupo de picos reveló una explosión en una cantera o construcción próxima. En otro registro apareció un frente frío representado por un conjunto continuado pero casual de picos. En dos ocasiones al menos, Lehman ha conseguido mostrar los efectos de los huracanes en el Atlántico.

Sin embargo, es probable que el lector centre su interés en conseguir registrar ondas sísmicas. Estas ondas son de cuatro tipos principales (existen otros tipos menos importantes y subtipos que no mencionaré). Dos de ellas se propagan a través de la tierra, pero de forma diferente y a distinta velocidad. Una es la onda transversal, llamada *S*. El material que participa en el movimiento de la onda oscila perpendicularmente a la dirección del movimiento de la misma. El término "transversal" significa que transmite ese sentido de oscilación. La mayoría de las ondas que conocemos son transversales; por ejemplo, las de la cuerda de una guitarra o de la superficie del agua.

El otro tipo de ondas que se extiende por la tierra son las longitudinales designadas *P*. Aquí el material que participa en el movimiento de la onda oscila hacia adelante y hacia atrás en la dirección en que se está moviendo la onda. El término "longitudinal" se usa aquí para indicar el hecho de que las oscilaciones son paralelas a la dirección de la onda.

Los otros dos tipos generales de ondas sísmicas se propagan a lo largo de la superficie de la tierra. Con un tipo, las denominadas ondas Love (*L*), el material que participa en el movimiento de la onda oscila transversalmente a la dirección del movimiento de la misma y horizontalmente con relación a la superficie de la tierra. Con el otro tipo, llamado ondas de Rayleigh (*R*), el material se mueve en cursos elípticos perpendiculares a la superficie y alineados en torno a la dirección de movimiento de la onda.

Si llegaran todas estas ondas a un sismógrafo al mismo tiempo, el aficionado no sería capaz de distinguirlas, pero no llegan a la vez porque o bien toman diferentes rutas o viajan a distintas velocidades. Las primeras en llegar son las ondas *P*; éstas y las ondas *S* toman un rumbo directo a través de la tierra, viajando las *P* más deprisa que las *S*. Después llegan las *S*, seguidas de los dos tipos de ondas superficiales, cuyo retraso se debe a que viajan sobre un largo recorrido por la superficie de la tierra.

El tiempo que transcurre entre la llegada de las *P* y las *S* se puede usar para determinar la distancia entre el observador y el centro del terremoto. En referencias sismográficas tales como el libro de Markus Bath citado en la bibliografía de este mismo número se encuentran tablas relativas a la diferencia en los tiempos de llegada y las distancias correspondientes. La distancia se cita con frecuencia en términos de la distancia angular entre el epicentro del terremoto y el sismógrafo medida desde el centro de la Tierra. Por ejemplo, si la diferencia en tiempos de llegada es unos 7 minutos 11 segundos, el epicentro estará a unos 50 grados del observador.

Con un instrumento simple se puede suponer sólo la dirección. La determinación aproximada de la misma puede hacerse con dos instrumentos orientados perpendicularmente entre sí (un sismógrafo siguiendo el eje norte-sur y el otro el este-oeste). La determinación precisa de la distancia, dirección y profundidad del epicentro se obtiene sólo cuando registran el hecho varias estaciones de sismógrafos muy separadas entre sí. La diferencia en tiempos de llegada de las ondas a las estaciones proporciona el dato de dónde se halla el epicentro.

Pero no sólo se pueden registrar ruidos de fondo y terremotos. Las explosiones nucleares subterráneas también envían ondas sísmicas a través de la tierra y a lo largo de su superficie. Lehman ha registrado cierto número de ellas que ocurrieron en Nevada, siete exactamente, en el primer trimestre de 1976.

Lehman me ha enviado varios ejemplos de datos de terremotos, dos de los cuales son especialmente llamativos. Uno es un registro del terremoto que sucedió en Las Galápagos el 29 de marzo de 1976 a las 0539:33 UTC. El movimiento tenía un epicentro poco profundo, una magnitud de 6,7 y aparentemente resultó del corrimiento a lo largo del borde entre las placas tectónicas de Cocos y Nazca. Seis minutos y 57 segundos después de haber empezado el terremoto, las ondas *P* alcanzaron el registro de Lehman. Tenían períodos que oscilaban de uno a tres segundos y fueron registradas durante más de cinco minutos. Las ondas *S* tardaron 12 minutos y 28 segundos en llegar, ya que tenían períodos más largos y mayores amplitudes que las *P*. Las ondas superficiales tardaron unos 15 minutos en llegar y tenían períodos desde 12 a 15 segundos. Las oscilaciones continuaron durante casi una hora. La diferencia en los tiempos de llegada de las ondas *P* y *S* (cinco minutos 31 segundos) indicó que el epicentro estaba a unos 3930 kilómetros del aparato de Lehman.

Libros

Fundamentación científica de la ecología, mapas geológicos y una teoría general del universo

Antonio Escarré, Alfredo García González y Luis Mas

PERSPECTIVAS DE LA TEORIA ECOLOGICA, por Ramón Margalef, traducción de la versión inglesa (1968) de Rosa Miracle; 110 págs. y 12 figuras. Editorial Blume, Barcelona 1979. Después de diez años de su publicación, cifra que excede bastante el retraso medio con que en nuestro país se aborda la traducción de lo que se consideran obras interesantes, y después también de su traducción al japonés en 1972, aparece ahora la versión castellana de uno de los libros de ecología más importantes de la década de los sesenta, y del que es autor el profesor catalán Ramón Margalef. Para intentar cuantificar el interés y la repercusión de esta obra, se puede decir que figura citada en casi el 5 por ciento de los artículos publicados en la revista americana *Ecology* en los últimos seis años.

El libro recoge cuatro conferencias dadas por Margalef en la Universidad de Chicago en 1966, cada una de las cuales constituye un capítulo. El primero de ellos es una visión cibernética del ecosistema; el segundo trata de la sucesión ecológica y la explotación de la naturaleza por parte del hombre; el tercero está dedicado a los ecosistemas pelágicos y el capítulo final presenta la evolución como un proceso que tiene lugar en el seno de los ecosistemas.

Margalef es un científico extraordinariamente imaginativo. Poseedor de una envidiable formación como naturalista, ha trabajado en el campo de la ciencia positiva, pero sin dejarse constreñir por un planteamiento ultraexperimentalista, y conservando toda la agudeza interpretativa de un biólogo de campo, ha dedicado buena parte de su vida a la búqueda de unos principios generales de la ecología que hagan de ella una ciencia respetable, fácilmente encajable en un contexto científico general.

Problemas derivados de la extraordinaria complejidad de la biosfera, como son la evaluación de las semejanzas existentes entre porciones de ecosistemas tan arbitrarias cuales las taxocenosis, no gozan de excesiva reputación como quehacer científico entre los investigadores experimentales que, curiosamente, son ahora más benevolentes al haberse generalizado el uso de modelos taximétricos multivariantes. La búsqueda de respetabilidad estaría pues justificada, pero no ha sido sin embargo la causa primordial en el desarrollo del pensamiento científico de Margalef, que más bien se ha preocupado por contemplar la ecología desde las generalizaciones conseguidas en otras ramas de la ciencia. Es precisamente en esa interfase entre diferentes ramas del saber donde su aportación científica resulta más fecunda.

El libro que me ocupa no es una obra sencilla, y quizá no sea demasiado adecuada para todos los que se balancean indistintamente entre los conceptos de ecología, ecologismo y calidad de vida. Se trata de un libro de teoría ecológica y que precisamente supera la visión casuística de la ecología que amenazaba con describir cada una de las interacciones entre organismo y medio, y cada uno de los tipos de ecosistema que constituyen la biosfera, antes de buscar alguna regularidad en el funcionamiento de la naturaleza. El interés por la "microscopía ecológica" del primer enfoque es sustituido aquí por el de la "macroscopía ecológica"; lo importante son los parámetros que permiten comparar un bosque con un lago o un cultivo de algas.

Entre los aspectos más sugestivos del libro debo destacar la presentación que se hace del acoplamiento de subsistemas con distinto grado de organización, del que podría deucirse que la asimetría es un principio básico en el funcionamiento de los sistemas biológicos: la mayor predictibilidad de los sistemas con elevado nivel de organización se consigue gracias al aporte energético de los sistemas menos maduros que, de esta forma, pierden opciones en su autoorganización. Esto equivale a decir que la regla más general en la relación funcional entre subsistemas biológicos es la de la explotación de unos por otros.

La idea medular de la obra es, sin duda, la sucesión ecológica que aparece como un proceso vital direccional, para el que resulta posible encontrar parámetros macroscópicos del ecosistema que presenten una variación regular en su transcurso: el cociente entre producción primaria y biomasa, el espectro de diversidad y la proporción de biomasas con diferentes tasas respiratorias son características que, por ejemplo, cumplen esta condición. El efecto del hombre sobre el resto de la naturaleza se puede catalogar también como una explotación. La tendencia de los ecosistemas a alcanzar cotas de organización elevadas se ve interrumpida por la acción humana que busca acoplarse con subsistemas con mayor flujo energético y, por decirlo de alguna manera, más "dóciles". La irrupción del hombre en la biosfera viene perfectamente ilustrada por la importante ampliación del canal etológico o cultural, en relación con los canales genético y ecológico de transmisión de información en el ecosistema. Esa transmisión por fuera del canal genético, de lo que se ha aprendido por propia experiencia, es la vía más reciente para la proyección de información hacia el futuro.

Actualmente se pueden ya observar los primeros síntomas de la aparición de un circuito recurrente negativo, que tiende a la defensa de la naturaleza frente a la acción del hombre, y la labor educativa se prevé como de vital importancia para conseguir la futura armonía entre el hombre y su medio. (A.E.)

MAPAS GEOLOGICOS: EXPLICACION E INTERPRETACION, por J.A. Martínez-Alvarez; 245 páginas y 157 figuras y cuadros desplegables. Editorial Paraninfo, Madrid 1979. En una época en que las publicaciones geológicas se presentan hacia direcciones muy especializadas, encaminadas principalmente a lectores con un grado de instrucción geológica elevado, es grato encontrarse con un trabajo muy práctico, en el que se representa de una manera sencilla lo que es la base de todo trabajo de geología: el mapa geológico.

Como es norma en el autor de esta obra, plasma su contenido de una manera general en las construcciones gráficas, haciéndolas base de toda su exposición. Desarrolla con gran amplitud y claridad la técnica de la representación tridimensional, forma realista de plasmar una representación geológica para ofrecerla más asequible a la captación del lector, principalmente cuando éste se encuentra en los períodos de iniciación a la geología. La obra consta de tres apartados, distintos en cuanto su concepción, si bien encadenados por un hilo conductor evolutivo hacia la profundización constante que va desde las ideas básicas

que todo lector de mapas geológicos debe poseer hasta la interpretación objetiva que nos ofrece la geología aplicada.

La primera parte se nos presenta de una manera escueta, como manual detallado de la diversidad de conceptos usados habitualmente en la confección de un mapa geológico. Expone una clara síntesis de los diversos procesos geológicos que han dado lugar a las configuraciones morfológicas que se hallan representadas en la corteza terrestre.

En la segunda parte, de carácter eminentemente técnico, el autor desarrolla de una manera amplia y objetiva las construcciones auxiliares que sobre el mapa geológico deben realizarse a fin de obtener los máximos datos posibles para una posterior interpretación. Hace una clara separación entre los datos propios localizables y los extrapolados. Describe las cualidades representativas de un corte geológico, construcción que generalmente se hace más interpretativa de lo que la escala del mapa permite.

La tercera parte, titulada "Lectura interpretativa del mapa geológico", está basada en los resultados provenientes de una recta interpretación entre la conjugación de los dos primeros apartados y la amplia experiencia adquirida en el desarrollo de la geología, ejemplarizada por los procesos geológicos propios de la Región Asturiana. (A. G.)

EL UNIVERSO. SU PRINCIPIO Y SU FIN, por Lloyd Motz. Editorial Antoni Bosch, Barcelona 1979. Los libros de divulgación presentan diferentes niveles de precisión y profundidad que se adaptan a los distintos públicos a los que van destinados. Se trata en este caso de un libro dirigido a personas que sólo tengan una preparación muy débil en física y química, incluso inferior a lo que se exige en Enseñanza Media. La dificultad de conseguir a este nivel un buen libro no es grano de anís. En particular, si la obra ha de cubrir *toda* la historia del universo. El principal problema consiste en saber distinguir lo que hay que justificar de lo que sólo hay que describir y, además, que el lector sea consciente de cuándo ocurre lo uno y lo otro.

La capacidad del autor para vencer este obstáculo es indudable y casi siempre distingue los resultados probados de aquellos todavía dudosos, si bien se inclina claramente hacia una teoría global del universo de un modo tan entusiasta que el lector debe sentirse forzado a aceptarla, aunque se le diga que no está todavía probada y sólo sea una conjetura, en plena concordancia con el título de la colección.

Comentemos algunos de los detalles del "argumento" del libro, "argumento" porque el libro se presenta como una apasionante historia con un planteo decidido, un largo desarrollo y un final casi feliz. Es característica de estas obras de divulgación la dificultad que encuentran en encajar, dentro de un cuadro elemental, una de las cuatro fuerzas de la naturaleza, la llamada fuerza o interacción débil que interviene en tantos artículos de *Investigación y Ciencia* sobre partículas elementales. Aquí tampoco se consigue y, después de una brevísima definición, no vuelve a hablarse de ella, y si acaso, queda englobada con la interacción fuerte en lo que el autor llama la fuerza nuclear.

Las teorías cosmológicas vienen bien explicadas, aunque siempre con el acento en la teoría de la gran explosión, evidentemente la que goza de mayor aceptación en nuestros días. La cosmología estacionaria, que tuvo su época de parcial adhesión en las décadas 1940 y 1950, no se basa en las ecuaciones de la teoría de Einstein sobre la relatividad general, sino únicamente en el llamado principio cosmológico perfecto, que incluye la perfecta homogeneidad e isotropía del universo en espacio y en tiempo. Si sólo incluimos el espacio, se deducen para el espacio-tiempo las llamadas métricas de Robertson-Walker, y añadiendo el carácter estacionario para el tiempo, queda completamente determinada la métrica del universo estacionario. Sin admitir esta última hipótesis, para acabar de determinar las métricas de Robertson-Walker, son necesarias las ecuaciones de la teoría de la relatividad general que conducen a los modelos de Friedman que, a su vez, incluyen los modelos de gran explosión.

A partir de las observaciones, se pretende hoy medir la constante de Hubble, el parámetro de deceleración, otro parámetro relacionado con la densidad del universo y la constante cosmológica y, finalmente, el valor o signo de la curvatura del universo. Para ello, tal como se explica en el libro tenemos tres grupos de observaciones: (*a*) Los desplazamientos hacia el rojo de las radiaciones recibidas en la Tierra, procedentes de otras estrellas o galaxias; (*b*) la medida de la intensidad e isotropía de la radiación de fondo que se mide en la Tierra y que parece corresponder ahora a la de un cuerpo negro que estaría a la temperatura de tres grados Kelvin y que sería el resto actual de la explosión inicial; y (*c*) abundancia en el universo de ciertos elementos químicos, con una posible medición de la densidad total.

El valor de la constante de Hubble varía según los observadores, y va de 50 a 90 kilómetros por segundos por megaparsec (km. seg^{-1}. $megaparsec^{-1}$), valores bastante más pequeños que los que dio Hubble. El margen en que puede variar es crítico, ya que de momento permite toda clase de modelos; e incluso los valores altos encajan con una constante cosmológica distinta de cero, rehabilitando así el siempre demasiado señalado error de Einstein, que no fue tal error sino una conjetura.

No es fácil medir el parámetro de deceleración, pues sus efectos son pequeños y se requiere una precisión de aparatos hoy en día casi inalcanzable. Se hace a partir de una expresión no lineal en los desplazamientos hacia el rojo, bien comprobada en su parte lineal utilizando galaxias cercanas. En la parte no lineal interviene el parámetro buscado. Los valores van desde −1,27 a 1,2, incluyendo el efecto de evolución porque la constante de Hubble depende de la edad del universo. Todos los modelos son posibles e incluso cabe que el parámetro de deceleración sea de aceleración si prevalecen los valores negativos.

El parámetro relacionado con la densidad y la constante cosmología nos indicaría, si fuera nulo, que también lo sería la constante cosmológica. Si le sumamos el parámetro de deceleración, y contando la densidad de las galaxias, un límite para el valor inferior sería 0,03 + 0,01, pero faltaría añadir la materia intergaláctica de muy difícil medición. A partir de teorías sobre los primeros instantes del universo y la producción de deuterio, se tiene un límite superior de 0,04. Existen otras observaciones que dan otros valores, pero, en general, tal como indica claramente en este caso el autor, todos los resultados actuales ofrecen valores demasiado pequeños si preferimos defender un universo en expansión que se frena.

La descripción de la formación de galaxias y estrellas y su evolución está muy lograda, y, como consecuencia de ello, la formación de los planetas. Al explicar la ley de Titius-Bode sobre colocación de las órbitas planetarias se insinúa una justificación o deducción moderna de ella, pero no se explicita; más adelante, al hablar del mundo cuántico, se dice que con otras condiciones iniciales las órbitas serían otras. Ello no tiene en cuenta la interacción entre planetas, pues recientemente se ha llegado al siguiente resultado: Se introducen en una calculadora datos del Sol y de los planetas con modificaciones, sino grandes sí sensibles, en cuanto a condiciones iniciales y masas, y se hacen evolucionar según las leyes de Newton. El resultado numérico es que las órbitas tienden siempre a las actuales, con lo que no sería tan aleatoria la distribución entre dichas órbitas planetarias. También den-

tro del movimiento de los planetas, cabe resaltar que no existe ninguna prueba rigurosa de la estabilidad de su movimiento cuando incluimos interacciones planeta-planeta.

Muy interesante, y bien descrito, es el capítulo de Tierra-Luna y las mareas. Queda claro que no importa tanto la fuerza de atracción en sí cuanto las variaciones de valor y dirección para las distintas partes de un sólido rígido, como la Tierra y la Luna. Esto mismo sería causa, de no poder observar un campo gravitatorio de gran intensidad y con simetría esférica demasiado cerca del centro, ya que la estructura sólida del observador humano o aparato correspondiente no resistiría la diferencia de fuerza atractiva sobre sus distintas partes y se desintegraría.

Al hablar del final de la vida de las estrellas, después de describir las formas más corrientes, el autor insiste en los modelos más catastróficos o violentos: las estrellas de neutrones y los agujeros negros. Las primeras están bien comprobadas y clasificadas, mientras que los segundos presentan hoy muchas dificultades teóricas y de observación. Para compensar el entusiasmo del autor, vamos a insistir en estas dificultades, sin que ello prejuzgue nada sobre el entusiasmo del comentarista y basándonos únicamente en datos observados y cálculos deducidos con total precisión.

Una de las propiedades o axiomas de la teoría de la relatividad general es el principio de covariancia, que afirma que las leyes físicas deben expresarse igual para cualquier observador. Este principio es de una sencillez y belleza sin igual y elimina los famosos y discutidos sistemas inerciales de las leyes de Newton y de la teoría de la relatividad especial. La consecuencia práctica de dicho principio es que las leyes físicas deben expresarse a través de campos tensoriales, siendo en este caso los campos de tensores, al revés de en teorías anteriores, verdaderas y completas aplicaciones multilineales del espacio-tiempo en sí mismo. La ley se plantea de un modo general y, luego, cada observación utilizará el sistema de coordenadas que más le convenga sin prácticamente otra condición que la de que localmente el sistema de coordenadas sea hiperbólico; por ejemplo, las famosas ecuaciones de Einstein de la relatividad general afirman que el tensor de Einstein (es el de Ricci más un término proporcional a la métrica) es igual al tensor de impulsión-energía de los campos tensoriales no gravitatorios.

Esta generalidad lleva consigo, sin embargo, una gran dificultad: uno debe saber si el sistema de coordenadas que elige es el que se adapta a las medidas que le dan sus aparatos, y, si las leyes no se verifican siempre, puede quedarse con la duda de que obedezca a esta falta de correspondencia (como el número de sistemas que puede escoger es infinito, difícil será eliminar la duda). Lo cual se complica todavía más si añadimos la posibilidad de que un sistema de coordenadas se aproxime localmente al que se adapta a nuestros aparatos de medida y que, sin embargo, no se adapte globalmente a ellos.

La mayoría de las propiedades de los agujeros negros se basan en la utilización de las coordenadas de Schwarzschild o análogas, que incluyen una singularidad, la llamada singularidad de Schwarzschild, que no es más que el horizonte de sucesos del agujero negro esférico. Existen, y esto es crucial en una discusión sobre si existen o no y de cómo son los agujeros negros, sistemas de coordenadas para los que no hay tal singularidad. Se han propuesto modelos globales de variedad espacio-tiempo en que no existe nada detrás del horizonte de sucesos.

Lo que sí constituye un resultado invariante es la existencia del horizonte de sucesos, pero no lo es el tiempo infinito que se dice tardaríamos en apreciar el colapso ni el carácter verdaderamente singular del comportamiento dentro del agujero negro, que tiende a reducir el volumen a cero. En estos casos de campo gravitatorio fuerte, hay que poner en duda la validez de las leyes conocidas de la gravedad, incluidas las de la relatividad general, y considerar, tal como indica el autor, el aspecto cuántico de las mismas, en fase de estudio intenso, que podría cambiar por entero las consecuencias. Por último, habría que añadir que no hay ningún astrofísico, incluidos los partidarios de los agujeros negros, que admita que el Sol sea candidato a un futuro agujero negro debido al valor de su masa. Se está estudiando la posibilidad de que la galaxia M87 tenga un núcleo de gran masa, lo que unido a su oscuridad, la haría un firme candidato a un agujero negro de los gigantes.

El optimismo final del autor lleva aparejado unas consideraciones sobre la belleza y unidad de las leyes que gobiernan el universo muy en la línea de las reflexiones de Einstein, reflejo de esta proclividad que tenemos a creer en la comprensibilidad de todo lo que nos rodea, si bien a veces uno puede dudar sobre si lo que nos dice el instinto se corresponde o no con la realidad. Todas estas sugerencias y otras muchas que pueden despertarse son un fruto excelente que debe siempre derivarse de la lectura de un buen libro de investigación. (L.M.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LOS EFECTOS INMEDIATOS Y RETARDADOS DE LA GUERRA NUCLEAR

LONG-TERM WORLDWIDE EFFECTS OF MULTIPLE NUCLEAR-WEAPONS DETONATIONS. National Academy of Sciences, Washington, D.C., 1975.

THE EFFECTS OF NUCLEAR WEAPONS. Dirigido por Samuel Glasstone y Philip Dolan. U.S. Departments of Defense and Energy, Government Printing Office, 1977.

SOVIET CIVIL DEFENSE. Informe del Director de la Oficina de Inteligencia, Bureau of Public Affairs, Office of Public Communications; julio, 1978.

EL VIRUS DE EPSTEIN-BARR

THE ONCOGENICITY OF EPSTEIN-BARR VIRUS. G. Miller en *The Journal of Infectious Diseases,* vol. 130, n.º 2, págs. 187-205; agosto, 1974.

THE EPSTEIN-BARR VIRUS AND NEOPLASIA. George Klein en *The New England Journal of Medicine,* vol. 293, n.º 26, págs. 1353-1357; 25 de diciembre de 1975.

RECENT PROGRESS IN EPSTEIN-BARR VIRUS RESEARCH. M. A. Epstein y B. G. Achong. en *Annual Review of Microbiology,* vol. 31, págs. 421-445; 1977.

THE IMMUNOLOGICAL APPROACH TO STUDY OF POSSIBLY VIRUS-INDUCED HUMAN MALIGNANCIES USING THE EPSTEIN-BARR VIRUS AS EXAMPLE. W. Henle y G. Henle en *Progress in Experimental Tumor Research,* vol. 21: *Viruses and Human Cancer,* dirigido por Yohei Ito. S. Karger, 1978.

EL PARSEC CENTRAL DE LA GALAXIA

INFRARED OBSERVATIONS OF THE GALACTIC CENTER, I: NATURE OF THE COMPACT SOURCES. E. E. Becklin, K. Matthews, G. Neugebauer y S. P. Willner en *The Astrophysical Journal,* vol. 219, n.º 1, parte 1, págs. 121-128; 1 de enero de 1978.

THE INFRARED EMISSION OF THE GALACTIC CENTER. G. H. Rieke, C. M. Telesco y D. A. Harper en *The Astrophysical Journal,* vol. 220, n.º 2, parte 1, págs. 556-567; 1 de marzo de 1978.

LOS ORIGENES DEL PRIMER AVION MOTOPROPULSADO CON PILOTO A BORDO

THE PAPERS OF WILBUR AND ORVILLE WRIGHT. Dirigido por Marvin W. McFarland. McGraw-Hill Book Company, Inc., 1953.

SIR GEORGE CAYLEY, INVENTOR OF THE AEROPLANE. John L. Pritchard. Horizon Press, 1962.

THE INVENTION OF THE AEROPLANE 1799-1909. Charles H. Gibbs-Smith. Faber and Faber, 1966.

WILBUR AND ORVILLE WRIGHT: A CHRONOLOGY COMMEMORATING THE HUNDREDTH ANNIVERSARY OF THE BIRTH OF ORVILLE WRIGHT, AUGUST 19, 1871. Arthur G. Renstrom. U.S. Library of Congress, Science and Technology Division, 1975.

COMPARTIMENTOS EN EL DESARROLLO DE LOS ANIMALES

COMPARTMENTS AND POLYCLONES IN INSECT DEVELOPMENT. F. H. C. Crick y P. A. Lawrence en *Science,* vol. 189, n.º 4200, págs. 340-347; 1 de agosto de 1975.

COMPARTMENTS IN THE WING OF *DROSOPHILA:* A STUDY OF THE *ENGRAILED* GENE. P. A. Lawrence y G. Morata en *Developmental Biology,* vol. 50, n.º 2, págs. 321-337; junio, 1976.

HOMOEOTIC GENES, COMPARTMENTS AND CELL DETERMINATION IN *DROSOPHILA.* G. Morata y Peter A. Lawrence en *Nature,* vol. 265, n.º 5591, págs. 211-216; 20 de enero de 1977.

HOMOEOTIC AND ATAVIC MUTATIONS IN INSECTS. A. García-Bellido en *American Zoologist,* vol. 17, n.º 3, págs. 613-629; verano, 1977.

EL MODELO EN BOLSA DEL CONFINAMIENTO DE LOS QUARKS

MASSES AND OTHER PARAMETERS OF THE LIGHT HADRONS. T. DeGrand, R. L. Jaffe, K. Johnson y J. Kiskis en *Physical Review, D.,* vol. 12, n.º 7, págs. 2060-2076; octubre, 1975.

STRINGLIKE SOLUTIONS OF THE BAG MODEL. K. Johnson y C. B. Thorn en *Physical Review D.,* vol. 13, n.º 7, págs. 1934-1939; 1 de abril de 1976.

QUARK CONFINEMENT R. L. Jaffe en *Nature,* vol. 268, n.º 5617, págs 201-208; 21 de julio de 1977.

LOS ANIMALES DE BURGESS SHALE

THE ENIGMATIC ANIMAL *OPABINIA REGALIS,* MIDDLE CAMBRIAN, BURGESS SHALE, BRITISH COLUMBIA. H. B. Whittington en *Philosophical Transactions of the Royal Society of London,* serie B, vol. 271, págs. 1-43; 26 de junio de 1975.

FOSSIL PRIAPULID WORMS. Simon Conway Morris en *Special Papers in Paleontology,* n.º 20. The Paleontological Association of London; diciembre, 1977.

THE BURGESS SHALE (MIDDLE CAMBRIAN) FAUNA. Simon Conway Morris en *Annual Review of Ecology and Systematics,* en prensa.

FREGE

ESTUDIO DE LOS FUNDAMENTOS DE LA ARITMETICA DE FREGE. (Apéndice de Frege. Fundamentos de la Aritmética.) Cl. Imbert. Ed. Laia, 1972. Versión de Ulises Moulines.

EL DESARROLLO DE LA LOGICA. W. y M. Kneale. Ed. Tecnos, 1972. Tradn. de Javier Muguerza.

SENTIDO Y REFERENCIA EN LA LOGICA DE GOTTLOB FREGE. Christian Thiel. Ed. Tecnos, 1972. Tradn. de José Sanmartín.

JUEGOS MATEMATICOS

GÖDEL'S PROOF. Ernest Nagel y James R. Newman. New York University Press. 1958.

THE BACH READER: A LIFE OF JOHANN SEBASTIAN BACH IN LETTERS AND DOCUMENTS. Dirigido por Hans T. David y Arthur Mendel. W. W. Norton & Co., 1966.

SEMANTIC INFORMATION PROCESSING. Dirigido por Marvin L. Minsky. The MIT Press, 1968.

TALLER Y LABORATORIO

ELEMENTARY SEISMOLOGY. Charles F. Richter. W. H. Freeman and Co., 1958.

EARTHQUAKE INTERPRETATIONS. Ruth S. Simon. Colorado School of Mines, 1969.

AN INTRODUCTION TO SEISMOLOGY. Markus Bath. John Wiley & Sons, Inc., 1973.